Probing cosmology with weak lensing selected clusters I: Halo approach and all-sky simulations

Masato Shirasaki[1][*], Takashi Hamana[1] and Naoki Yoshida[2,3,4]

[1]*National Astronomical Observatory of Japan, Mitaka, Tokyo 181-8588, Japan*
[2]*Department of Physics, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan*
[3]*Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study,*
The University of Tokyo, Kashiwa, Chiba 277-8583, Japan
[4]*CREST, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan*

ABSTRACT

We explore a variety of statistics of clusters selected with cosmic shear measurement by utilizing both analytic models and large numerical simulations. We first develop a halo model to predict the abundance and the clustering of weak lensing selected clusters. Observational effects such as galaxy shape noise are included in our model. We then generate realistic mock weak lensing catalogs to test the accuracy of our analytic model. To this end, we perform full-sky ray-tracing simulations that allow us to have multiple realizations of a large continuous area. We model the masked regions on the sky using the actual positions of bright stars, and generate 200 mock weak lensing catalogs with sky coverage of ∼1000 squared degrees. We show that our theoretical model agrees well with the ensemble average of statistics and their covariances calculated directly from the mock catalogues. With a typical selection threshold, ignoring shape noise correction causes overestimation of the clustering of weak lensing selected clusters with a level of about 10%, and shape noise correction boosts the cluster abundance by a factor of a few. We calculate the cross-covariances using the halo model with accounting for the effective reduction of the survey area due to masks. The covariance of the cosmic shear auto power spectrum is affected by the mode-coupling effect that originates from sky masking. Our model and the results can be readily used for cosmological analysis with ongoing and future weak lensing surveys.

Key words: gravitational lensing: weak — cosmological parameters cosmology: theory — large-scale structure in the universe

1 INTRODUCTION

The accelerating expansion of the universe is now established by an array of astronomical observations such as Type Ia supernovae (e.g., Betoule et al. 2014), measurement of baryon acoustic oscillations in galaxy surveys (e.g., Beutler et al. 2011; Blake et al. 2011; Anderson et al. 2014), anisotropies of the cosmic microwave background (CMB) (e.g., Hinshaw et al. 2013; Planck Collaboration et al. 2014), large-scale galaxy distribution (e.g., Sánchez et al. 2012; Beutler et al. 2014), and weak gravitational lensing (e.g., Kilbinger et al. 2013). In order to realize the cosmic acceleration within the theory of general relativity, an exotic form of energy needs to be postulated to dominate in the present-day universe. There is another possibility to explain the cosmic acceleration without dark energy, e.g., modified gravity theory. Modified gravity models do not assume an unknown energy but change essentially the basic equation of gravitational action. Observationally, measurement of the growth of matter density fluctuations will help us to distinguish the models including the Einstein gravity with dark energy, because the modification of gravity induce characteristic clustering patterns in matter density distribution.

[*] E-mail: masato.shirasaki@nao.ac.jp

Gravitational lensing is a powerful probe of the matter distribution in the universe. Small image distortions of distant galaxies are caused by intervening mass distribution. Small distortion caused by the large-scale structure of the universe is called cosmic shear. It contains, in principle, rich information on the matter distribution at small and large scales and the evolution over time. Image distortion induced by gravitational lensing is, however, very small in general. Therefore, we need statistical analyses of the cosmic shear signal by sampling a large number of distant galaxies in order to extract cosmological information from gravitational lensing. The conventional statistic of cosmic shear is two-point correlation function or its Fourier-counterpart, power spectrum. If the cosmic shear field obeys a Gaussian distribution, the two-point statistics suffice to describe all the information of cosmic shear. However, this is not the case in reality because cosmic shear has non-Gaussian information caused by non-linear gravitational growth (Sato et al. 2009). In order to extract the full information content, it is desirable to use other statistical quantities that probes nonlinear structure of length scale of ∼10 Mpc or less. Clusters of galaxies are one the most reliable objects for this purpose.

The number count of clusters is expected to be highly sensitive to growth of matter density perturbations (Lilje 1992), whereas the spatial correlation of the position of clusters and cosmic shear provides the information on the matter density profile as well as clustering of clusters (e.g., Oguri et al. 2012; Okabe et al. 2013; Covone et al. 2014). Fortunately, cosmic shear itself provides an efficient way of locating galaxies of clusters. Cluster finding methods with cosmic shear are based on reconstruction of matter density distribution over an area of sky (Hamana, Takada & Yoshida 2004; Hennawi & Spergel 2005; Maturi et al. 2005; Marian et al. 2012). A reconstructed mass density map can be used to identify high density regions as "peaks" that are mostly caused by massive collapsed objects such as clusters of galaxies (Miyazaki et al. 2007; Schirmer et al. 2007; Shan et al. 2012). The unique advantage of weak lensing among various techniques is that it does not rely on uncertain physical state of the baryonic component in clusters. There have been a number of studies that investigate cosmological information in number counts of weak lensing selected clusters (Maturi et al. 2010; Kratochvil, Haiman & May 2010; Dietrich & Hartlap 2010; Yang et al. 2011; Hilbert et al. 2012). Recently, Marian et al. (2013) combined other statistics beyond the abundance of weak lensing selected clusters. The authors conclude that correlation analysis of weak lensing selected clusters allow one to derive tight constraints on cosmological parameters.

In the present paper, we study in detail the properties of a class of statistics of weak lensing selected clusters. Our study is aimed at being applied to real observations such as Subaru Hyper-Suprime-Cam Survey. It is well-known that the intrinsic ellipticities of source galaxies induce noise to lensing shear maps. The so-called shape noise causes typically false detection of clusters with cosmic shear measurement. We develop theoretical framework to model the correspondence of underlying dark matter halos and weak lensing selected clusters in presence of shape noise. Sky masking causes another important observational effect on statistical analyses of weak lensing selected clusters (Liu et al. 2014b). Since reconstructed mass density is usually defined by local cosmic shear signals, boundaries of masked regions would make the reconstruction inaccurate. In order to realize the realistic situation in galaxy imaging surveys, we perform gravitational lensing simulations on curved full-sky. We then utilize these simulations to create two hundreds of mock weak lensing catalogs with the proposed sky coverage in ongoing Hyper Suprime-Cam (HSC) survey[1]. For a large set of mock HSC surveys, we generate reconstructed mass map and identify the local maxima on each map as an indicator of weak lensing selected clusters. These simulations enable us to study the statistical property of weak lensing selected clusters in presence of shape noise and masked region. We are also able to examine our theoretical model through the large set of realistic mock weak lensing observations.

The rest of the paper is organized as follows. In Section 2, we describe the methodology to search clusters with cosmic shear measurement. There, we present the statistical property of weak lensing selected clusters and the theoretical model of statistics of interest. In Section 3, we use a large set of N-body simulations to perform full-sky lensing simulations and create mock weak lensing maps incorporated with the information of HSC surveys. In Section 4, we provide the result of our measurement of statistical quantities over a set of full-sky and masked sky simulations. We also compare the simulation results and our theoretical models in detail. Conclusions and discussions are summarized in Section 5.

2   WEAK LENSING

We summarize the basics of weak gravitational lensing effect in this section. We also describe the finder algorithm of galaxy clusters with weak lensing measurement.

---

[1]   http://www.naoj.org/Projects/HSC/index.html

2.1 Basics

When considering the observed position of a source object as $\boldsymbol{\theta}$ and the true position as $\boldsymbol{\beta}$, one can characterize the distortion of image of a source object by the following 2D matrix:

$$A_{ij} = \frac{\partial \beta^i}{\partial \theta^j} \equiv \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}, \tag{1}$$

where κ is convergence and γ is shear.

One can relate each component of A_{ij} to the second derivative of the gravitational potential as follows (Bartelmann & Schneider 2001; Munshi et al. 2008);

$$\begin{align} A_{ij} &= \delta_{ij} - \phi_{ij}, \tag{2} \\ \phi_{ij} &= \frac{2}{c^2} \int_0^\chi \mathrm{d}\chi' g(\chi, \chi') \partial_i \partial_j \Phi(\chi'), \tag{3} \end{align}$$

$$g(\chi, \chi') = \frac{r(\chi - \chi')r(\chi')}{r(\chi)}, \tag{4}$$

where χ is the comoving distance and $r(\chi)$ represents the comoving angular diameter distance. Gravitational potential Φ can be related to matter density perturbation δ according to Poisson equation. Therefore, convergence can be expressed as the weighted integral of δ along the line of sight;

$$\kappa = \frac{3}{2} \left(\frac{H_0}{c} \right)^2 \Omega_{\mathrm{m0}} \int_0^\chi \mathrm{d}\chi' g(\chi, \chi') \frac{\delta}{a}. \tag{5}$$

2.2 Cluster finding

Weak lensing provides a physical method to reconstruct the projected matter density field. The conventional technique for reconstruction is based on the smoothed map of cosmic shear. Let us first define the smoothed convergence field as

$$\mathcal{K}(\boldsymbol{\theta}) = \int \mathrm{d}^2\theta' \, \kappa(\boldsymbol{\theta} - \boldsymbol{\theta}')U(\boldsymbol{\theta}'), \tag{6}$$

where U is the filter function to be specified below. Although we can calculate the same quantity by smoothing the shear field γ (e.g., Shirasaki & Yoshida 2014), we use Eq. (6) for simplicity in the following.

Various functional form of U are proposed in literature (e.g., Hamana, Takada & Yoshida 2004; Hennawi & Spergel 2005; Maturi et al. 2010; Marian et al. 2012). We consider the Gaussian filter[2]

$$U(\theta) = \frac{1}{\pi \theta_G^2} \exp\left(-\frac{\theta^2}{\theta_G^2} \right). \tag{7}$$

With this filter, we can easily model the statistical properties of the contaminant of a smoothed $\mathcal{K}$ map, called shape noise. The noise in a $\mathcal{K}$ map would follow the Gaussian distribution when one can use a sufficient large number of source galaxies and when source galaxies are oriented randomly. The Gaussian properties of the noise makes it easy to model the lensing peak statistics, as will be shown in the following.

We denote the shape noise contribution to a smoothed lensing map by $\mathcal{N}$. For a given smoothing scale θ_G, correlation function of the shape noise after Gaussian smoothing is given by (van Waerbeke 2000)

$$\langle \mathcal{N}(\boldsymbol{\theta})\mathcal{N}(\boldsymbol{\theta}') \rangle = \frac{\sigma_\gamma^2}{4\pi n_{\mathrm{gal}} \theta_G^2} \exp\left[-\frac{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2}{2\theta_G^2} \right], \tag{8}$$

where σ_γ is the rms of the intrinsic ellipticity of sources and n_{gal} represents the number density of source galaxies. One can derive the power spectrum of noise convergence field $\mathcal{N}$ by Fourier transforming of Eq. (8);

$$P_\mathcal{N}(\ell) = \frac{\sigma_\gamma^2}{2n_{\mathrm{gal}}} \exp\left[-\frac{1}{2}\theta_G^2 \ell^2 \right]. \tag{9}$$

Using Eq. (9), we define the moment of $\mathcal{N}$ as

$$\sigma_{\mathrm{noise},i} = \left(\int \frac{\mathrm{d}^2\boldsymbol{\ell}}{(2\pi)^2} \, \ell^{2i} P_\mathcal{N}(\ell) \right)^{1/2}. \tag{10}$$

---

[2] In practice, the filter function should be compensated in order to remove an undetermined constant convergence (Schneider 1996). In Appendix B, we examine compensated Gaussian filters when searching for weak-lensing clusters. There, we show that our model can be suitably modified for the case of compensated Gaussian filters.

In a smoothed lensing map, peaks with high signal-to-noise ratio $\nu = \mathcal{K}/\sigma_{\mathrm{noise},0}$ are likely associated with cluster of galaxies (e.g., Hamana, Takada & Yoshida 2004). We first locate high peaks on a $\mathcal{K}$ map and then relate each peak with an isolated massive halo along the line of sight. We assume the following universal density profile of dark matter halos (Navarro, Frenk & White 1997):

$$\rho_h(r) = \frac{\rho_s}{(r/r_s)\,(1 + r/r_s)^2},\tag{11}$$

where r_s and ρ_s represent the scale radius and the scale density, respectively. The parameters r_s and ρ_s can be essentially convolved into one parameter, the concentration $c_{\mathrm{vir}}(M, z)$, by the use of two halo mass relations; namely, $M = 4\pi r_{\mathrm{vir}}^3 \Delta_{\mathrm{vir}}(z)\rho_{\mathrm{crit}}(z)/3$, where r_{vir} is the virial radius corresponding to the overdensity criterion $\Delta_{\mathrm{vir}}(z)$ as shown in, e.g., Navarro, Frenk & White (1997), and $M = \int dV \,\rho_h(r_s, \rho_s)$ with the integral performed out to r_{vir}. In this paper, we adopt the functional form of the concentration parameter in Duffy et al. (2008),

$$c_{\mathrm{vir}}(M, z) = 5.72 \left(\frac{M}{10^{14}h^{-1}M_\odot}\right)^{-0.081}(1 + z)^{-0.71}.\tag{12}$$

The corresponding convergence can be calculated as in Hamana, Takada & Yoshida (2004),

$$\kappa_h(R) = \frac{2\rho_s r_s f(R/r_s)}{\Sigma_{\mathrm{crit}}},\tag{13}$$

where R represents the perpendicular proper distance from the center of halo and $f(x)$ is

$$f(x) = \begin{cases} -\frac{\sqrt{c_{\mathrm{vir}}^2 - x^2}}{(1-x^2)(1+c_{\mathrm{vir}})} + \frac{1}{(1-x^2)^{3/2}}\mathrm{arccosh}\left[\frac{x^2 + c_{\mathrm{vir}}}{x(1+c_{\mathrm{vir}})}\right] & (x < 1), \\ \frac{\sqrt{c_{\mathrm{vir}}^2 - 1}}{3(1+c_{\mathrm{vir}})}\left(1 + \frac{1}{1+c_{\mathrm{vir}}}\right) & (x = 1), \\ -\frac{\sqrt{c_{\mathrm{vir}}^2 - x^2}}{(1-x^2)(1+c_{\mathrm{vir}})} - \frac{1}{(1-x^2)^{3/2}}\arccos\left[\frac{x^2 + c_{\mathrm{vir}}}{x(1+c_{\mathrm{vir}})}\right] & (1 < x \leq c_{\mathrm{vir}}), \\ 0 & (x > c_{\mathrm{vir}}). \end{cases}\tag{14}$$

In Eq. (13), Σ_{crit} is defined by the following relation

$$\Sigma_{\mathrm{crit}} = \frac{c^2}{4\pi G}\frac{D_{\mathrm{s}}}{D_{\mathrm{l}}D_{\mathrm{ls}}},\tag{15}$$

where D_{s}, D_{l}, and D_{ls} are the angular diameter distance to the source, to the lens, and between the source and the lens, respectively.

In order to predict the peak height in $\mathcal{K}$ map, we need to take the following effects into account: (i) the offset between the position of a peak and the center of the corresponding halo and (ii) the modulation of peak height due to the shape noise. Fan, Shan & Liu (2010) have studied these two effects using numerical simulations and analytic approach. Let us first work on the simple assumption that the peak position is set to be the halo center. The peak height in absence of shape noise is given by

$$\mathcal{K}_{\mathrm{peak},h} = \int \mathrm{d}^2\theta\, U(\theta; \theta_G)\kappa_h(\theta).\tag{16}$$

The actual peak height on a noisy $\mathcal{K}$ map is not given by Eq. (16), but it obeys a probability distribution (Fan, Shan & Liu 2010). The probability distribution function for a given $\mathcal{K}_{\mathrm{peak},h}$ is calculated by

$$\mathrm{Prob}(\mathcal{K}_{\mathrm{peak,obs}}|\mathcal{K}_{\mathrm{peak},h}) = \frac{n_{\mathrm{peak,N}}(\mathcal{K}_{\mathrm{peak,obs}}|\mathcal{K}_{\mathrm{peak},h})}{\int n_{\mathrm{peak,N}}(\mathcal{K}'_{\mathrm{peak,obs}}|\mathcal{K}_{\mathrm{peak},h})\mathrm{d}\mathcal{K}'_{\mathrm{peak,obs}}},\tag{17}$$

where $\mathcal{K}_{\mathrm{peak,obs}}$ is the measured peak height and $n_{\mathrm{peak,N}}$ is defined as the expected number density of peaks with the measured peak height of $\mathcal{K}_{\mathrm{peak,obs}}$ when the halo contribution $\mathcal{K}_{\mathrm{peak},h}$ is known in advance. For derivation of $n_{\mathrm{peak,N}}$, we decompose the observed peak height $\mathcal{K}_{\mathrm{peak,obs}}$ into three components:

$$\mathcal{K}_{\mathrm{peak,obs}} = \mathcal{N} + \mathcal{K}_{\mathrm{LSS}} + \mathcal{K}_{\mathrm{peak},h},\tag{18}$$

where $\mathcal{N}$ is the noise convergence field caused by shape noise, $\mathcal{K}_{\mathrm{LSS}}$ and $\mathcal{K}_{\mathrm{peak},h}$ represent the convergence field due to large-scale structure and foreground halos, respectively. Note that $\mathcal{K}_{\mathrm{peak},h}$ is a known quantity to derive $n_{\mathrm{peak,N}}$. We aim at determining the relationship between $\mathcal{K}_{\mathrm{peak,obs}}$ and a given $\mathcal{K}_{\mathrm{peak},h}(z, M)$.

Following Fan, Shan & Liu (2010), we assume that the noise field $\mathcal{N}$ is given by a Gaussian distribution with the power spectrum of Eq. (9). If $\mathcal{K}_{\mathrm{LSS}}$ is a Gaussian random field, at the position of peaks, the *total* noise field (i.e. $\mathcal{N} + \mathcal{K}_{\mathrm{LSS}}$) obeys the probability distribution function of Gaussian peaks. Also, we can calculate the contribution from the (known) corresponding halo once the difference between the peak position and the halo center is specified. We assume that the peak position is at the center of the corresponding halo. although the equality does not hold in general. We have checked that the assumption is indeed reasonable for peaks with high signal-to-noise ratio in the case of $\theta_G \sim 2\,\mathrm{arcmin}$, $\sigma_\gamma = 0.4$, and $n_{\mathrm{gal}} \gtrsim 10\,\mathrm{arcmin}^{-2}$.
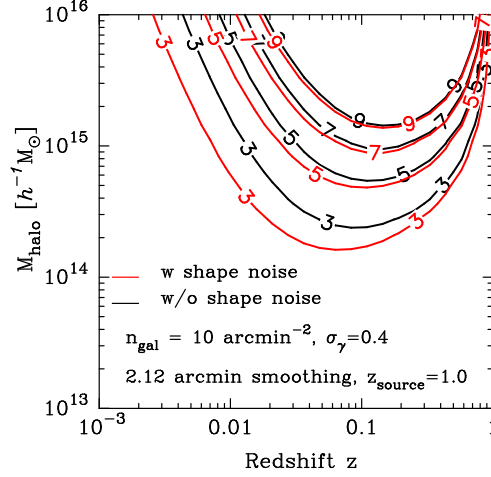
Figure 1. The effective selection function of galaxy clusters. Black line shows the expected convergence signal caused by an isolated dark matter halo. Red line shows the modulated peak height on a noisy smoothed convergence map (see the text for the definition). We plot each line in units of the variance of shape noise $\sigma_{\mathrm{noise},0} = 0.017$.

Therefore, we set $n_{\mathrm{peak,N}}$ to be the number density of peaks for Gaussian field $\mathcal{N}+\mathcal{K}_{\mathrm{LSS}}$. Using the relation of $\mathcal{N}+\mathcal{K}_{\mathrm{LSS}} = \mathcal{K}_{\mathrm{peak,obs}} - \mathcal{K}_{\mathrm{peak},h}$, we can obtain the number density $n_{\mathrm{peak,N}}$ as (see, Fan, Shan & Liu (2010) for details)

$$
n_{\mathrm{peak,N}}(\mathcal{K}_{\mathrm{peak,obs}}|\mathcal{K}_{\mathrm{peak},h}) = \frac{1}{2\pi\theta_*^2}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\mathcal{K}_{\mathrm{peak,obs}} - \mathcal{K}_{\mathrm{peak},h}}{\sigma_0}\right)^2\right]\int \frac{\mathrm{d}x_N}{[2\pi(1-\gamma_N^2)]^{1/2}}\, F(x_N|\mathcal{K}_{\mathrm{peak},h})
$$

$$
\times \exp\left\{-\frac{1}{2(1-\gamma_N^2)}\left[x_N + \left(\mathcal{K}_{\mathrm{peak},h}^{11} + \mathcal{K}_{\mathrm{peak},h}^{22}\right)/\sigma_2 - \gamma_N\left(\mathcal{K}_{\mathrm{peak,obs}} - \mathcal{K}_{\mathrm{peak},h}\right)/\sigma_0\right]^2\right\}, \quad (19)
$$

where σ_i is the ith moment of $\mathcal{N}+\mathcal{K}_{\mathrm{LSS}}$, $\theta_*^2 = 2(\sigma_1/\sigma_2)^2$, $\gamma_N = \sigma_1^2/(\sigma_0\sigma_2)$ and $\mathcal{K}_{\mathrm{peak},h}^{ii}$ denotes the second derivative of $\mathcal{K}_{\mathrm{peak},h}$ with respect to θ_i at the halo centre. Here, $F(x_N|\mathcal{K}_{\mathrm{peak},h})$ is calculated as follows:

$$
F(x_N|\mathcal{K}_{\mathrm{peak},h}) = \exp\left[-\left(\frac{\mathcal{K}_{\mathrm{peak},h}^{11} - \mathcal{K}_{\mathrm{peak},h}^{22}}{\sigma_2}\right)^2\right]\int_0^{1/2}\mathrm{d}e_N\, 8(x_N^2 e_N)x_N^2(1-4e_N^2)\exp\left(-4x_N^2 e_N^2\right)
$$

$$
\times\int_0^\pi \frac{\mathrm{d}\theta_N}{\pi}\exp\left[-4x_N e_N \cos 2\theta_N\frac{\mathcal{K}_{\mathrm{peak},h}^{11} - \mathcal{K}_{\mathrm{peak},h}^{22}}{\sigma_2}\right]. \quad (20)
$$

In Eqs. (19) and (20), x_N and e_N are given by

$$
x_N = \frac{\lambda_{N1} + \lambda_{N2}}{\sigma_2}, \quad e_N = \frac{\lambda_{N1} + \lambda_{N2}}{2\sigma_2 x_N}, \quad (21)
$$

where λ_{Ni} represents the ith diagonal component of the second derivative tensor of $\mathcal{N}+\mathcal{K}_{\mathrm{LSS}}$.

We adopt $\sigma_\gamma = 0.4$, $n_{\mathrm{gal}} = 10\,\mathrm{arcmin}^{-2}$ and the source redshift is set to be $z_{\mathrm{source}} = 1$. These are the typical values for ground-based galaxy imaging surveys (e.g., Heymans et al. 2012). Also, we employ a Gaussian smoothing with the full width at half maximum of 5 arcmin. This corresponds to $\theta_G = 5/\sqrt{8\ln 2} = 2.12$ arcmin and thus to $\sigma_{\mathrm{noise},0} \simeq 0.017$. Although the smoothing scale adopted here is slightly larger than that in previous works (see, e.g., Hamana, Takada & Yoshida 2004) by a factor of about two, the noise level on the smoothed convergence map is similar. Gaussian smoothing with $\theta_G \sim 2$ arcmin with the actual data set is already examined in Shan et al. (2012).

Let us examine the effect of shape noise on weak lensing peaks. For this purpose, we define the mean modulation of peak height in a noisy $\mathcal{K}$ map as follows:

$$
\bar{\mathcal{K}}_{\mathrm{peak,obs}}(z, M) = \int \mathrm{d}\mathcal{K}\,\mathcal{K}\,\mathrm{Prob}(\mathcal{K}|\mathcal{K}_{\mathrm{peak},h}(z, M)). \quad (22)
$$

Figure 1 shows the comparison with $\bar{\mathcal{K}}_{\mathrm{peak,obs}}(z, M)$ and $\mathcal{K}_{\mathrm{peak},h}(z, M)$ for a given dark matter halo with mass of M at redshift z. In this figure, red line shows the contour of $\bar{\mathcal{K}}_{\mathrm{peak,obs}}(z, M)$ in units of $\sigma_{\mathrm{noise},0}$, whereas black line indicates the contour of $\mathcal{K}_{\mathrm{peak},h}(z, M)$. In a noisy $\mathcal{K}$ map, the shape noise modulates the height of peaks and the number of peaks slightly increases. We have tested the validity of our model against numerical simulations. The result is shown in Appendix A.

As shown in Figure 1, the Gaussian smoothing of ~ 2 arcmin are effective to search for clusters with mass of $\sim 10^{14}h^{-1}M_\odot$

at $z \sim 0.1 - 0.2$. The selection of mass and redshift is basically determined by the typical angular size of dark matter halos of interest (e.g., Hamana, Takada & Yoshida 2004). Naively, it is expected that lower redshift clusters are detected with larger smoothing scales. In order to verify this expectation, we have studied the statistical properties of lensing peaks when adopting a Gaussian smoothing with the full width at half maximum of 15 arcmin (corresponding to $\theta_G \sim 6.4$ arcmin). In this case, we do not find one-to-one correspondence between selected peaks and halos, and thus our analytic model does not work. This is likely caused by the so-called projection effect; the effective redshift of lensing in the case of $z_{\rm source} = 1$ is $0.1 - 0.5$ whereas we attempt to search for clusters at lower redshift $z < 0.1$. We argue that our model is valid when the smoothing scale is set to be $1 - 2$ arcmin, corresponding to the typical angular size of dark matter halos with mass of $\sim 10^{14} h^{-1} M_\odot$ at $z \sim 0.1 - 0.5$.

2.3 Statistics

We consider a set of statistics derived from weak lensing measurement. In order to extract cosmological information from the number and the distribution of massive dark matter halos, we utilize peaks on a smoothed lensing map as described in Section 2.2.

Convergence power spectrum

First, we consider the power spectrum of convergence. Under the flat sky approximation, the Fourier transform of convergence field is defined by

$$\kappa(\boldsymbol{\theta}) = \int \frac{\mathrm{d}^2 \ell}{(2\pi)^2} e^{i\boldsymbol{\ell} \cdot \boldsymbol{\theta}} \tilde{\kappa}(\boldsymbol{\ell}). \tag{23}$$

The power spectrum of convergence field $P_{\kappa\kappa}$ is defined by

$$\langle \tilde{\kappa}(\boldsymbol{\ell}_1) \tilde{\kappa}(\boldsymbol{\ell}_2) \rangle = (2\pi)^2 \delta_D^{(2)}(\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2) P_{\kappa\kappa}(\ell_1), \tag{24}$$

where $\delta_D^{(2)}(\boldsymbol{\ell})$ is the Dirac delta function. By using Limber approximation[3] (Limber 1954; Kaiser 1992) and Eq. (5), one can calculate the convergence power spectrum as follows:

$$P_{\kappa\kappa}(\ell) = \int_0^{\chi_s} \mathrm{d}\chi \frac{W_\kappa(\chi)^2}{r(\chi)^2} P_\delta \left(k = \frac{\ell}{r(\chi)}, z(\chi) \right), \tag{25}$$

where $P_\delta(k)$ is the three dimensional matter power spectrum, χ_s is comoving distance to source galaxies and $W_\kappa(\chi)$ is the lensing weight function defined as

$$W_\kappa(\chi) = \frac{3}{2} \left(\frac{H_0}{c} \right)^2 \Omega_{\rm m0} \frac{r(\chi_s - \chi) r(\chi)}{r(\chi_s)} (1 + z(\chi)). \tag{26}$$

The non-linear gravitational growth of density fluctuations significantly affects the amplitude of convergence power spectrum at the angular scales less than 1 degree (Jain, Seljak & White 2000; Hilbert et al. 2009; Sato et al. 2009). Typical weak lensing surveys aim at measuring the angular scales larger than a few arcmin corresponding to a few Mpc. This leads weak lensing can be one of the most powerful probes for constraints on dark matter distribution at Mpc scale. Therefore, accurate theoretical prediction of non-linear matter power spectrum is essential for deriving cosmological constraints from weak lensing power spectrum. In order to predict the non-linear evolution of $P_\delta(k)$ for standard ΛCDM universe, a numerical approach based on N-body simulations has given steady results over the past few decades (Peacock & Dodds 1996; Smith et al. 2003; Heitmann et al. 2010; Takahashi et al. 2012). We adopt the most recent model of non-linear $P_\delta(k)$ of Takahashi et al. (2012).

Convergence peak count

The number count of massive clusters is sensitive to various cosmological parameters such as the equation of state of dark energy (e.g., Allen, Evrard & Mantz 2011). In this section, we use peak counts as a cosmological probe. We locate the local maxima in a smoothed lensing map and associate each identified peak with a massive dark matter halo along the same line of sight.

In practice, one can select a lensing peak by its peak height. We define the signal-to-noise ratio of a peak by $\nu =$

---

[3] The validity of Limber approximation have been discussed in e.g., Jeong, Komatsu & Jain (2009). The typical accuracy of Limber approximaion is of a level of $\lesssim 1\%$ for $\ell > 10$.

$\mathcal{K}_{\mathrm{peak,obs}}/\sigma_{\mathrm{noise},0}$. For a given threshold ν_{thre}, one can predict the surface number density of peaks with $\nu > \nu_{\mathrm{thre}}$ as follows (e.g., Hamana, Takada & Yoshida 2004):

$$N_{\mathrm{peak}}(\nu_{\mathrm{thre}}) = \int \mathrm{d}z\,\mathrm{d}M\, \frac{\mathrm{d}^2 V}{\mathrm{d}z\mathrm{d}\Omega} \frac{\mathrm{d}n}{\mathrm{d}M}(z,M) \int_{\nu_{\mathrm{thre}}\sigma_{\mathrm{noise},0}}^{\infty} \mathrm{d}\mathcal{K}_{\mathrm{peak,obs}}\, \mathrm{Prob}(\mathcal{K}_{\mathrm{peak,obs}}|\mathcal{K}_{\mathrm{peak},h}(z,M)), \tag{27}$$

where $\mathrm{d}n/\mathrm{d}M$ represents the mass function of dark matter halo and the volume element is expressed as $\mathrm{d}^2V/\mathrm{d}z\mathrm{d}\Omega = \chi^2/H(z)$ for a spatially flat universe. Here, $\mathrm{Prob}(\mathcal{K}_{\mathrm{peak,obs}}|\mathcal{K}_{\mathrm{peak},h})$ is given by Eq. (17). In the following, we adopt the model of halo mass function in Bhattacharya et al. (2011).

Convergence peak auto spectrum and cross spectrum

We next consider the auto-correlation function of peaks, and the peak-convergence cross correlation. Marian et al. (2013) study cosmological information obtained from the statistics using a large set of numerical simulations. They conclude that using the auto- and cross-correlation functions can improve the constraints on cosmological parameters when combined with the number of peaks. We develop an analytic halo model in order to predict the correlation function of peaks and cross correlation between peaks and convergence.

In the halo model, the number density field of weak lensing selected clusters is given by

$$
\begin{aligned}
n_{\mathrm{cl}}(\boldsymbol{x}) &= \sum_i \delta_D^{(3)}(\boldsymbol{x}-\boldsymbol{x}_i) S(z_i, M_i)\\
&= \sum_i \int \mathrm{d}M\, S(z,M)\, \delta_D(M-M_i) \int \mathrm{d}^3\boldsymbol{x}'\, \delta_D^{(3)}(\boldsymbol{x}'-\boldsymbol{x}_i)\delta_D^{(3)}(\boldsymbol{x}-\boldsymbol{x}'),
\end{aligned}
\tag{28}
$$

where $S(z,M)$ represents the selection function (e.g., Takada & Bridle 2007). For clusters identified as lensing peaks, $S(z,M)$ is expressed as

$$S(z,M|\nu_{\mathrm{thre}}) = \int_{\nu_{\mathrm{thre}}\sigma_{\mathrm{noise},0}}^{\infty} \mathrm{d}\mathcal{K}_{\mathrm{peak,obs}}\, \mathrm{Prob}(\mathcal{K}_{\mathrm{peak,obs}}|\mathcal{K}_{\mathrm{peak},h}(z,M)). \tag{29}$$

Also, underlying matter density field can be approximated as

$$
\begin{aligned}
\rho_m(\boldsymbol{x}) &= \sum_i \rho_h(\boldsymbol{x}-\boldsymbol{x}_i|z,M)\\
&= \sum_i \int \mathrm{d}M\, M\, \delta_D(M-M_i) \int \mathrm{d}^3\boldsymbol{x}'\, \delta_D^{(3)}(\boldsymbol{x}'-\boldsymbol{x}_i)u_m(\boldsymbol{x}-\boldsymbol{x}'|z,M),
\end{aligned}
\tag{30}
$$

where ρ_h is the density profile of a massive halo given by Eq. (11), and $\rho_h(r|z,M) = Mu_m(r|z,M)$.

In order to derive the auto power spectrum of weak lensing selected clusters for a given threshold ν_{thre}, we first consider the auto power spectrum of n_{cl}. The two point correlation function of n_{cl} is given by

$$
\begin{aligned}
\bar{n}_{\mathrm{cl}}^2 \xi_{cc}(\boldsymbol{x}_1 - \boldsymbol{x}_2) &\equiv \langle n_{\mathrm{cl}}(\boldsymbol{x}_1) n_{\mathrm{cl}}(\boldsymbol{x}_1)\rangle - \bar{n}_{\mathrm{cl}}^2\\
&= \langle \sum_i S^2(M_i)\delta_D^{(3)}(\boldsymbol{x}_1-\boldsymbol{x}_i)\delta_D^{(3)}(\boldsymbol{x}_2-\boldsymbol{x}_i)\rangle + \langle \sum_{i,j|\,i\neq j} S(M_i)S(M_j)\delta_D^{(3)}(\boldsymbol{x}_1-\boldsymbol{x}_i)\delta_D^{(3)}(\boldsymbol{x}_2-\boldsymbol{x}_j)\rangle\\
&= \langle \sum_i \int \mathrm{d}M \int \mathrm{d}^3 y\, S^2(M)\delta_D(M-M_i)\delta_D^{(3)}(\boldsymbol{x}_1-\boldsymbol{y})\delta_D^{(3)}(\boldsymbol{x}_2-\boldsymbol{y})\delta_D^{(3)}(\boldsymbol{y}-\boldsymbol{x}_i)\rangle\\
&\quad + \langle \sum_{i,j|\,i\neq j} \int \mathrm{d}M \int \mathrm{d}^3 y\, S(M)\delta_D(M-M_i)\delta_D^{(3)}(\boldsymbol{x}_1-\boldsymbol{y})\delta_D^{(3)}(\boldsymbol{y}-\boldsymbol{x}_i)\\
&\quad\quad\quad\quad \times \int \mathrm{d}M' \int \mathrm{d}^3 y'\, S(M)\delta_D(M'-M_j)\delta_D^{(3)}(\boldsymbol{x}_2-\boldsymbol{y}')\delta_D^{(3)}(\boldsymbol{y}'-\boldsymbol{x}_j)\rangle\\
&= \int \mathrm{d}M\, \frac{\mathrm{d}n}{\mathrm{d}M} S^2(M) \int \mathrm{d}^3 y\, \delta_D^{(3)}(\boldsymbol{x}_1-\boldsymbol{y})\delta_D^{(3)}(\boldsymbol{x}_2-\boldsymbol{y})\\
&\quad + \int \mathrm{d}M\, \frac{\mathrm{d}n}{\mathrm{d}M} S(M) \int \mathrm{d}^3 y\, \delta_D^{(3)}(\boldsymbol{x}_1-\boldsymbol{y}) \int \mathrm{d}M'\, \frac{\mathrm{d}n}{\mathrm{d}M'} S(M') \int \mathrm{d}^3 y'\, \delta_D^{(3)}(\boldsymbol{x}_2-\boldsymbol{y}')\xi_{hh}(\boldsymbol{y}-\boldsymbol{y}';M,M')\\
&= \int \mathrm{d}M\, \frac{\mathrm{d}n}{\mathrm{d}M} S^2(M) \int \mathrm{d}^3 y\, \delta_D^{(3)}(\boldsymbol{x}_1-\boldsymbol{x}_2) + \int \mathrm{d}M\, \frac{\mathrm{d}n}{\mathrm{d}M} S(M) \int \mathrm{d}M'\, \frac{\mathrm{d}n}{\mathrm{d}M'} S(M')\xi_{hh}(\boldsymbol{x}_1-\boldsymbol{x}_2;M,M'),
\end{aligned}
\tag{31}
$$

where $\xi_{hh}(\boldsymbol{y}-\boldsymbol{y}';M,M')$ is the two point correlation function of dark matter haloes with mass of M and M'. In the above

calculation, we use the following relations as

$$\left\langle \sum_i \delta_D(M - M_i)\delta_D^{(3)}(\boldsymbol{x} - \boldsymbol{x}_i) \right\rangle \equiv \frac{\mathrm{d}n}{\mathrm{d}M}, \tag{32}$$

$$\left\langle \sum_{i,j \mid i \neq j} \delta_D(M - M_i)\delta_D(M' - M_j)\delta_D^{(3)}(\boldsymbol{x}_1 - \boldsymbol{x}_i)\delta_D^{(3)}(\boldsymbol{x}_2 - \boldsymbol{x}_j) \right\rangle \equiv \frac{\mathrm{d}n}{\mathrm{d}M}\frac{\mathrm{d}n}{\mathrm{d}M'}\xi_{hh}(\boldsymbol{x}_1 - \boldsymbol{x}_2'; M, M'). \tag{33}$$

We further approximate ξ_{hh} as $b_h(M)b_h(M')\xi_{\delta\delta}^L$, where b_h is the linear halo bias and $\xi_{\delta\delta}^L$ is the two point correlation function of linear matter density field. Then, we finally obtain the following equation:

$$\bar{n}_{\mathrm{cl}}^2 \xi_{cc}(\boldsymbol{x}_1 - \boldsymbol{x}_2) = \int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} S^2(M)\delta_D^{(3)}(\boldsymbol{x}_1 - \boldsymbol{x}_2) + \left[\int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} S(M)b_h(M) \right]^2 \xi_{\delta\delta}^L(\boldsymbol{x}_1 - \boldsymbol{x}_2). \tag{34}$$

The Fourier transform of Eq. (34) is the three-dimensional power spectrum of n_{cl}, which is expressed as

$$P_{cc}(k) = \int \mathrm{d}^3 r \, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \xi_{cc}(r). \tag{35}$$

The observed surface number density of weak lensing selected clusters (for a given threshold ν_{thre}) is thus given by

$$p(\boldsymbol{\theta}) = \frac{1}{N_{\mathrm{peak}}(\nu_{\mathrm{thre}})} \int \mathrm{d}\chi \frac{\mathrm{d}^2 V}{\mathrm{d}\chi\mathrm{d}\Omega} n_{\mathrm{cl}}(r(\chi)\boldsymbol{\theta}, z(\chi)), \tag{36}$$

where N_{peak} is defined by Eq. (27). Using the Limber approximation, we can derive the angular power spectrum of p as

$$P_{\mathrm{pp}}(\ell) = \int \mathrm{d}\chi \frac{1}{r(\chi)^2} \left(\frac{1}{N_{\mathrm{peak}}} \frac{\mathrm{d}^2 V}{\mathrm{d}\chi\mathrm{d}\Omega} \right)^2 \bar{n}_{\mathrm{cl}}^2 P_{cc}\left(k = \frac{\ell}{r(\chi)}, z(\chi) \right). \tag{37}$$

Except for the shot noise term in Eq. (42), we obtain

$$P_{\mathrm{pp}}(\ell) = \int \mathrm{d}\chi \frac{1}{r(\chi)^2} \left(\frac{1}{N_{\mathrm{peak}}} \frac{\mathrm{d}^2 V}{\mathrm{d}\chi\mathrm{d}\Omega} \right)^2 \left[\int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M}(z, M)S(z, M|\nu_{\mathrm{thre}})b_h(z, M) \right]^2 P_m^L\left(k = \frac{\ell}{r(\chi)}, z(\chi) \right), \tag{38}$$

where P_m^L is the linear matter power spectrum. Throughput this paper, we adopt the functional form of b_h proposed in Bhattacharya et al. (2011).

The similar derivation can be applied to the cross power spectrum of weak lensing selected clusters and lensing convergence. Let us first consider the three-dimensional cross correlation function of n_{cl} and ρ_m:

$$\bar{\rho}_m \xi_{c\delta}(\boldsymbol{x}_1 - \boldsymbol{x}_2) = \langle n_{\mathrm{cl}}(\boldsymbol{x}_1)\rho_m(\boldsymbol{x}_1) \rangle - \bar{n}_{\mathrm{cl}}\bar{\rho}_m$$

$$= \int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} S(M)M u_m(\boldsymbol{x}_1 - \boldsymbol{x}_2|M) + \left[\int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} S(M)b_h(M) \right] \bar{\rho}_m \xi_{\delta\delta}^L(\boldsymbol{x}_1 - \boldsymbol{x}_2). \tag{39}$$

Through the derivation of Eq. (39), we use the following fact that

$$\bar{\rho}_m = \int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} b_h(M)M \int \mathrm{d}^3 x' \, u_m(\boldsymbol{x} - \boldsymbol{x}'|M). \tag{40}$$

Then, we can obtain the angular cross power spectrum of p and lensing convergence κ by the similar calculation as Eq. (37). The cross power spectrum $P_{\mathrm{p}\kappa}$ is given by

$$P_{\mathrm{p}\kappa}(\ell) = \int \mathrm{d}\chi \frac{W_\kappa(\chi)}{r(\chi)^2} \left(\frac{1}{N_{\mathrm{peak}}} \frac{\mathrm{d}^2 V}{\mathrm{d}\chi\mathrm{d}\Omega} \right) P_{c\delta}\left(k = \frac{\ell}{r(\chi)}, z(\chi) \right), \tag{41}$$

where $P_{c\delta}(k)$ represents the three-dimensional cross power spectrum of n_{cl} and matter overdensity field δ, i.e.,

$$P_{c\delta}(k) = \int \mathrm{d}^3 r \, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \xi_{c\delta}(r). \tag{42}$$

Finally, the cross power spectrum between peaks and convergence is given by (also, see Oguri & Takada (2011))

$$P_{\mathrm{p}\kappa}(\ell) = P_{\mathrm{p}\kappa}^{1h}(\ell) + P_{\mathrm{p}\kappa}^{2h}(\ell), \tag{43}$$

$$P_{\mathrm{p}\kappa}^{1h}(\ell) = \int \mathrm{d}\chi \frac{W_\kappa(\chi)}{r(\chi)^2} \left(\frac{1}{N_{\mathrm{peak}}} \frac{\mathrm{d}^2 V}{\mathrm{d}\chi\mathrm{d}\Omega} \right) \int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} S(z, M|\nu_{\mathrm{thre}}) \left(\frac{M}{\bar{\rho}_m(z)} \right) \tilde{u}_m\left(k = \frac{\ell}{r(\chi)} \Big| z(\chi), M \right), \tag{44}$$

$$P_{\mathrm{p}\kappa}^{2h}(\ell) = \int \mathrm{d}\chi \frac{W_\kappa(\chi)}{r(\chi)^2} \left(\frac{1}{N_{\mathrm{peak}}} \frac{\mathrm{d}^2 V}{\mathrm{d}\chi\mathrm{d}\Omega} \right) \left[\int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M}(z, M)S(z, M|\nu_{\mathrm{thre}})b_h(z, M) \right] P_m^L\left(k = \frac{\ell}{r(\chi)}, z(\chi) \right), \tag{45}$$

where $\tilde{u}_m$ is the Fourier transform of $u_m(r|z, M)$.

2.4   Covariances between statistics

We summarize covariance matrices between statistics of interest in the following. As a first approximation, we can use the Gaussian covariance between three binned spectra $P_{\kappa\kappa}$, $P_{\mathrm{p}\kappa}$ and P_{pp} as,

$$\mathrm{Cov}[P_{XY}(\ell), P_{AB}(\ell')] = \frac{1}{(2\ell+1)\Delta\ell f_{\mathrm{sky}}}\left[P_{XA}^{\mathrm{obs}}(\ell)P_{YB}^{\mathrm{obs}}(\ell) + P_{XB}^{\mathrm{obs}}(\ell)P_{YA}^{\mathrm{obs}}(\ell)\right]\delta_{\ell\ell'}, \tag{46}$$

where $\Delta\ell$ is the width of binning in multipole and f_{sky} represents the observed sky fraction. The observed spectra $P_{XY}^{\mathrm{obs}}(\ell)$ are then defined by

$$P_{\kappa\kappa}^{\mathrm{obs}} = P_{\kappa\kappa} + \frac{\sigma_\gamma^2}{2n_{\mathrm{gal}}}, \; P_{\mathrm{pp}}^{\mathrm{obs}} = P_{\mathrm{pp}} + \frac{1}{N_{\mathrm{peak}}}, \; P_{\mathrm{p}\kappa}^{\mathrm{obs}} = P_{\mathrm{p}\kappa}. \tag{47}$$

The non-linear gravitational growth causes mode-coupling of the density fluctuations with different wavelengths. The mode-coupling then induces the correlation of weak lensing statistics between different multipoles, i.e., we can not use the Gaussian approximation to covariances (e.g., Sato et al. 2009). Modeling of the non-Gaussian covariances is still being developed (e.g., Cooray & Hu 2001; Takada & Jain 2004; Takada & Bridle 2007). We here present a theoretical model of non-Gaussian covariance matrices between $P_{\kappa\kappa}$, $P_{\mathrm{p}\kappa}$ and N_{peak}.

Let us consider the following set of four-point correlation functions in Fourier space:

$$\langle\tilde{\delta}_m(\boldsymbol{k}_1)\tilde{\delta}_m(\boldsymbol{k}_2)\tilde{\delta}_m(\boldsymbol{k}_3)\tilde{\delta}_m(\boldsymbol{k}_4)\rangle = (2\pi)^3\delta_D^{(3)}(\boldsymbol{k}_{1234})T_{\delta\delta\delta\delta}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \tag{48}$$

$$\langle\tilde{\delta}_{cl}(\boldsymbol{k}_1)\tilde{\delta}_m(\boldsymbol{k}_2)\tilde{\delta}_{cl}(\boldsymbol{k}_3)\tilde{\delta}_m(\boldsymbol{k}_4)\rangle = (2\pi)^3\delta_D^{(3)}(\boldsymbol{k}_{1234})T_{c\delta c\delta}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \tag{49}$$

$$\langle\tilde{\delta}_{cl}(\boldsymbol{k}_1)\tilde{\delta}_m(\boldsymbol{k}_2)\tilde{\delta}_m(\boldsymbol{k}_3)\tilde{\delta}_m(\boldsymbol{k}_4)\rangle = (2\pi)^3\delta_D^{(3)}(\boldsymbol{k}_{1234})T_{c\delta\delta\delta}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4), \tag{50}$$

where $\boldsymbol{k}_{ij\cdots n} = \boldsymbol{k}_i + \boldsymbol{k}_j + \cdots + \boldsymbol{k}_n$, δ_m and δ_{cl} represent over density of matter and weak lensing selected clusters, respectively. In the flat sky approximation, we can relate three tri-spectra ($T_{\delta\delta\delta\delta}, T_{c\delta c\delta}, T_{c\delta\delta\delta}$) with the non-Gaussian part of the covariance matrix of weak lensing statistics as follows:

$$\mathrm{Cov}\left[P_{\kappa\kappa}(\ell), P_{\kappa\kappa}(\ell')\right]_{\mathrm{NG}} = \frac{1}{4\pi}\int\frac{\mathrm{d}\phi}{2\pi}T_{\kappa\kappa\kappa\kappa}(\boldsymbol{\ell},-\boldsymbol{\ell},\boldsymbol{\ell}',-\boldsymbol{\ell}';\phi), \tag{51}$$

$$\mathrm{Cov}\left[P_{\mathrm{p}\kappa}(\ell), P_{\mathrm{p}\kappa}(\ell')\right]_{\mathrm{NG}} = \frac{1}{4\pi}\int\frac{\mathrm{d}\phi}{2\pi}T_{\mathrm{p}\kappa\mathrm{p}\kappa}(\boldsymbol{\ell},-\boldsymbol{\ell},\boldsymbol{\ell}',-\boldsymbol{\ell}';\phi), \tag{52}$$

$$\mathrm{Cov}\left[P_{\mathrm{p}\kappa}(\ell), P_{\kappa\kappa}(\ell')\right]_{\mathrm{NG}} = \frac{1}{4\pi}\int\frac{\mathrm{d}\phi}{2\pi}T_{\mathrm{p}\kappa\kappa\kappa}(\boldsymbol{\ell},-\boldsymbol{\ell},\boldsymbol{\ell}',-\boldsymbol{\ell}';\phi), \tag{53}$$

where ϕ is the angle between two vectors $\boldsymbol{\ell}$ and $\boldsymbol{\ell}'$. In practice, the integral over ϕ is often simplified as, e.g.,

$$\frac{1}{4\pi}\int\frac{\mathrm{d}\phi}{2\pi}T_{\kappa\kappa\kappa\kappa}(\boldsymbol{\ell},-\boldsymbol{\ell},\boldsymbol{\ell}',-\boldsymbol{\ell}';\phi) \simeq T_{\kappa\kappa\kappa\kappa}(\ell,\ell,\ell',\ell'), \tag{54}$$

and so on. With Limber approximation, $T_{\kappa\kappa\kappa\kappa}$, $T_{\mathrm{p}\kappa\mathrm{p}\kappa}$, and $T_{\mathrm{p}\kappa\kappa\kappa}$ are given by

$$T_{\kappa\kappa\kappa\kappa}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_3,\boldsymbol{\ell}_4) = \int_0^{\chi_s}\mathrm{d}\chi\frac{W_\kappa^4}{r(\chi)^6}T_{\delta\delta\delta\delta}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4;z(\chi)), \tag{55}$$

$$T_{\mathrm{p}\kappa\mathrm{p}\kappa}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_3,\boldsymbol{\ell}_4) = \int_0^{\chi_s}\mathrm{d}\chi\frac{W_\kappa^2}{r(\chi)^6}\left(\frac{1}{N_{\mathrm{peak}}}\frac{\mathrm{d}^2V}{\mathrm{d}\chi\mathrm{d}\Omega}\right)^2 T_{c\delta c\delta}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4;z(\chi)), \tag{56}$$

$$T_{\mathrm{p}\kappa\kappa\kappa}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_3,\boldsymbol{\ell}_4) = \int_0^{\chi_s}\mathrm{d}\chi\frac{W_\kappa^3}{r(\chi)^6}\left(\frac{1}{N_{\mathrm{peak}}}\frac{\mathrm{d}^2V}{\mathrm{d}\chi\mathrm{d}\Omega}\right) T_{c\delta\delta\delta}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4;z(\chi)), \tag{57}$$

where $k_i = l_i/\chi$, χ_s is the comoving distance to sources, the window function W_κ is given by Eq. (26) and N_{peak} is defined by Eq. (27).

Previous works show that the dominant contribution of the non-Gaussian covariance is the so-called one-halo term of the relevant tri-spectrum at $\ell \gtrsim 100$ (e.g., Sato et al. 2009). One halo term arises from the four point correlation between different modes $\boldsymbol{k}_i$ ($i = 1, 2, 3, 4$) in a single dark matter halo. Thus, one halo term of the underlying tri-spectra between δ_m and δ_c can be calculated as

$$T_{\delta\delta\delta\delta}^{1h}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4;z) = \int\mathrm{d}M\frac{\mathrm{d}n}{\mathrm{d}M}(z,M)\left(\frac{M}{\bar{\rho}_m(z)}\right)^4\tilde{u}_m(\boldsymbol{k}_1|z,M)\tilde{u}_m(\boldsymbol{k}_2|z,M)\tilde{u}_m(\boldsymbol{k}_3|z,M)\tilde{u}_m(\boldsymbol{k}_4|z,M), \tag{58}$$

$$T_{c\delta c\delta}^{1h}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4;z) = \int\mathrm{d}M\frac{\mathrm{d}n}{\mathrm{d}M}(z,M)\left(\frac{M}{\bar{\rho}_m(z)}\right)^2 S(z,M)\tilde{u}_m(\boldsymbol{k}_2|z,M)\tilde{u}_m(\boldsymbol{k}_4|z,M), \tag{59}$$

$$T_{c\delta\delta\delta}^{1h}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4;z) = \int\mathrm{d}M\frac{\mathrm{d}n}{\mathrm{d}M}(z,M)\left(\frac{M}{\bar{\rho}_m(z)}\right)^3 S(z,M)\tilde{u}_m(\boldsymbol{k}_2|z,M)\tilde{u}_m(\boldsymbol{k}_3|z,M)\tilde{u}_m(\boldsymbol{k}_4|z,M). \tag{60}$$

Another important contribution of covariance at degree scales or less is so-called halo sampling variance (HSV) (Sato et al.

2009; Kayo, Takada & Jain 2013). This term describes the mode-coupling between the measured Fourier modes and larger modes of length-scales comparable to survey volume. It is expected to be important when the number of massive haloes found in a finite region is correlated with the overall mass density fluctuation in the region (Hu & Kravtsov 2003). Following Kayo, Takada & Jain (2013), we model the HSV of weak lensing statistics as

$$
\begin{aligned}
\mathrm{Cov}\left[P_{\kappa\kappa}(\ell), P_{\kappa\kappa}(\ell')\right]_{\mathrm{HSV}} &= \int_0^{\chi_s} \mathrm{d}\chi \left(\frac{\mathrm{d}^2 V}{\mathrm{d}\chi \mathrm{d}\Omega}\right)^2 \left[\int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} b_h(M) |\tilde{\kappa}_h(\ell|M)|^2\right] \\
&\times \left[\int \mathrm{d}M' \frac{\mathrm{d}n}{\mathrm{d}M'} b_h(M') |\tilde{\kappa}_h(\ell'|M')|^2\right] \left[\int_0^\infty \frac{k\mathrm{d}k}{2\pi} P_m^L(k) |\tilde{W}(k\chi\Theta_{\mathrm{survey}})|^2\right],
\end{aligned} \tag{61}
$$

$$
\begin{aligned}
\mathrm{Cov}\left[P_{\mathrm{p}\kappa}(\ell), P_{\mathrm{p}\kappa}(\ell')\right]_{\mathrm{HSV}} &= \int_0^{\chi_s} \mathrm{d}\chi \left(\frac{\mathrm{d}^2 V}{\mathrm{d}\chi \mathrm{d}\Omega}\right)^2 \left[\int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} b_h(M) S(z,M) \tilde{\kappa}_h(\ell|M)\right] \\
&\times \left[\int \mathrm{d}M' \frac{\mathrm{d}n}{\mathrm{d}M'} b_h(M') S(z,M') \tilde{\kappa}_h(\ell'|M')\right] \left[\int_0^\infty \frac{k\mathrm{d}k}{2\pi} P_m^L(k) |\tilde{W}(k\chi\Theta_{\mathrm{survey}})|^2\right],
\end{aligned} \tag{62}
$$

$$
\begin{aligned}
\mathrm{Cov}\left[P_{\mathrm{p}\kappa}(\ell), P_{\kappa\kappa}(\ell')\right]_{\mathrm{HSV}} &= \int_0^{\chi_s} \mathrm{d}\chi \left(\frac{\mathrm{d}^2 V}{\mathrm{d}\chi \mathrm{d}\Omega}\right)^2 \left[\int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} b_h(M) S(z,M) \tilde{\kappa}_h(\ell|M)\right] \\
&\times \left[\int \mathrm{d}M' \frac{\mathrm{d}n}{\mathrm{d}M'} b_h(M') |\tilde{\kappa}_h(\ell'|M')|^2\right] \left[\int_0^\infty \frac{k\mathrm{d}k}{2\pi} P_m^L(k) |\tilde{W}(k\chi\Theta_{\mathrm{survey}})|^2\right],
\end{aligned} \tag{63}
$$

where κ_h is the Fourier transforming of Eq. (13). Here, $\tilde{W}(k\chi\Theta_{\mathrm{survey}})$ represents the window function of the survey region in Fourier space with Θ_{survey} denoting the squared root of the survey area. We use the circular function of $\tilde{W}(x) = J_1(x)/x$.

Cross covariance between the number count of weak lensing selected clusters and the lensing spectra can be naturally incorporated in the halo model (Takada & Bridle 2007; Takada & Spergel 2014).

The covariance of N_{peak} with two different thresholds is given by

$$
\begin{aligned}
\mathrm{Cov}[N_{\mathrm{peak}}(\nu_{\mathrm{thre},1}), N_{\mathrm{peak}}(\nu_{\mathrm{thre},2})] &= \int \mathrm{d}^2\theta \, W(\boldsymbol{\theta}) \int \mathrm{d}\theta' \, W(\boldsymbol{\theta}') \int \mathrm{d}\chi \frac{\mathrm{d}^2 V}{\mathrm{d}\chi \mathrm{d}\Omega} \int \mathrm{d}\chi' \frac{\mathrm{d}^2 V}{\mathrm{d}\chi' \mathrm{d}\Omega'} \\
&\times \left[n_{\mathrm{cl}}(\chi\boldsymbol{\theta}, z(\chi)|\nu_{\mathrm{thre},1}) n_{\mathrm{cl}}(\chi'\boldsymbol{\theta}', z(\chi')|\nu_{\mathrm{thre},2}) - \bar{n}_{\mathrm{cl}}(z(\chi)|\nu_{\mathrm{thre},1}) \bar{n}_{\mathrm{cl}}(z(\chi')|\nu_{\mathrm{thre},2})\right],
\end{aligned} \tag{64}
$$

where $W(\boldsymbol{\theta})$ represents the window function in real space. Performing the similar calculation as in Eq. (31), one can find that (see also the appendix in Takada & Bridle (2007))

$$
\begin{aligned}
\mathrm{Cov}[N_{\mathrm{peak}}(\nu_{\mathrm{thre},1}), N_{\mathrm{peak}}(\nu_{\mathrm{thre},2})] &= \delta_{12}\left\{\frac{N_{\mathrm{peak}}(\nu_{\mathrm{thre},1})}{4\pi f_{\mathrm{sky}}} + \int \mathrm{d}\chi \left(\frac{\mathrm{d}^2 V}{\mathrm{d}\chi \mathrm{d}\Omega}\right)^2 r(\chi)^{-2} \left[\int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} b_h(M) S(z,M|\nu_{\mathrm{thre},1})\right]^2\right. \\
&\left.\times \left[\int_0^\infty \frac{k\mathrm{d}k}{2\pi} P_m^L(k) |\tilde{W}(k\chi\Theta_{\mathrm{survey}})|^2\right]\right\},
\end{aligned} \tag{65}
$$

where f_{sky} is the sky fraction for survey of interest. In order to derive the cross covariance between N_{peak} and $P_{\kappa\kappa}$ or $P_{\mathrm{p}\kappa}$, we first define the estimator of power spectrum as

$$
P_{\kappa\kappa}^{\mathrm{est}}(\ell) = \frac{1}{4\pi f_{\mathrm{sky}} N_p(\ell)} \sum_{\boldsymbol{\ell}} \tilde{\kappa}(\boldsymbol{\ell}) \tilde{\kappa}(-\boldsymbol{\ell}), \quad P_{\mathrm{p}\kappa}^{\mathrm{est}}(\ell|\nu_{\mathrm{thre}}) = \frac{1}{4\pi f_{\mathrm{sky}} N_p(\ell)} \sum_{\boldsymbol{\ell}} \tilde{p}(\boldsymbol{\ell}|\nu_{\mathrm{thre}}) \tilde{\kappa}(-\boldsymbol{\ell}), \tag{66}
$$

where the summation is taken over all the Fourier modes in the range of $[\ell - \Delta\ell/2, \ell + \Delta\ell/2]$ and $\Delta\ell$ is the width of multipoles. Also, $N_p(\ell)$ represents the number of modes to estimate of power spectrum with the multipole of ℓ. With Eq. (66), we can express the cross covariance of N_{peak} and $P_{\kappa\kappa}$ as

$$
\begin{aligned}
\mathrm{Cov}[N_{\mathrm{peak}}(\nu_{\mathrm{thre}}), P_{\kappa\kappa}(\ell)] &= \frac{1}{4\pi f_{\mathrm{sky}} N_p(\ell)} \sum_{\boldsymbol{\ell}} \int \mathrm{d}^2\theta \, W(\boldsymbol{\theta}) \langle \tilde{\kappa}(\boldsymbol{\ell}) \tilde{\kappa}(-\boldsymbol{\ell}) [n_{\mathrm{cl}}(\chi\boldsymbol{\theta}, z(\chi)|\nu_{\mathrm{thre}}) - \bar{n}_{\mathrm{cl}}(z(\chi)|\nu_{\mathrm{thre}})]\rangle, \\
&= \frac{1}{4\pi f_{\mathrm{sky}} N_p(\ell)} \int \mathrm{d}^2\theta \, W(\boldsymbol{\theta}) N_{\mathrm{peak}}(\nu_{\mathrm{thre}}) \sum_{\boldsymbol{\ell}} \sum_{\boldsymbol{\ell}'} \langle \tilde{\kappa}(\boldsymbol{\ell}) \tilde{\kappa}(-\boldsymbol{\ell}) \tilde{p}(\boldsymbol{\ell}'|\nu_{\mathrm{thre}})\rangle e^{-i\boldsymbol{\ell}' \cdot \boldsymbol{\theta}},
\end{aligned} \tag{67}
$$

where we use Eq. (36) through the derivation. Similarly, the cross covariances of N_{peak} and $P_{\mathrm{p}\kappa}$ is given by

$$
\begin{aligned}
\mathrm{Cov}[N_{\mathrm{peak}}(\nu_{\mathrm{thre},1}), P_{\mathrm{p}\kappa}(\ell|\nu_{\mathrm{thre},2})] &= \frac{1}{4\pi f_{\mathrm{sky}} N_p(\ell)} \int \mathrm{d}^2\theta \, W(\boldsymbol{\theta}) N_{\mathrm{peak}}(\nu_{\mathrm{thre},2}) \\
&\times \sum_{\boldsymbol{\ell}} \sum_{\boldsymbol{\ell}'} \langle \tilde{p}(\boldsymbol{\ell}|\nu_{\mathrm{thre},1}) \tilde{\kappa}(-\boldsymbol{\ell}) \tilde{p}(\boldsymbol{\ell}'|\nu_{\mathrm{thre},2})\rangle e^{-i\boldsymbol{\ell}' \cdot \boldsymbol{\theta}}.
\end{aligned} \tag{68}
$$

Therefore, the cross covariances of N_{peak} and lensing spectra include the three point correlation of the relevant field p or κ.

$L_{\rm box}\,[h^{-1}{\rm Mpc}]$	$z_{\rm init}$	No. of sim.	output redshift
450	72	10	0.025, 0.076, 0.129
900	36	10	0.182, 0.237, 0.294
1350	24	10	0.352, 0.412, 0.475
1800	18	10	0.540, 0.607, 0.677
2250	15	10	0.751, 0.827, 0.901
2700	12	10	0.990, 1.077, 1.169

Table 1. Parameters used for N-body simulations. Each simulation was run with 1024^3 dark matter particles. The output redshift of each simulation corresponds to the comoving distance to the center of lens-shells. We adopted the standard ΛCDM model, which is consistent with WMAP nine-year results (Hinshaw et al. 2013).

We can thus summarize the relevant covariance are

$$\mathrm{Cov}[N_{\rm peak}(\nu_{\rm thre}), P_{\kappa\kappa}(\ell)] \;=\; \frac{1}{4\pi f_{\rm sky}} \int \mathrm{d}\chi\, \frac{\mathrm{d}^2 V}{\mathrm{d}\chi\mathrm{d}\Omega}\, \frac{W_\kappa^2}{r(\chi)^4}\, B_{c\delta\delta}(0, \ell/\chi, \ell/\chi; z(\chi)), \tag{69}$$

$$\mathrm{Cov}[N_{\rm peak}(\nu_{\rm thre,1}), P_{\rm p\kappa}(\ell|\nu_{\rm thre,2})] \;=\; \frac{1}{4\pi f_{\rm sky}} \int \mathrm{d}\chi\, \left(\frac{\mathrm{d}^2 V}{\mathrm{d}\chi\mathrm{d}\Omega}\right)^2 \frac{W_\kappa}{r(\chi)^4}\, B_{c(1)c(2)\delta}(0, \ell/\chi, \ell/\chi; z(\chi)), \tag{70}$$

where $B_{c\delta\delta}$ and $B_{c(1)c(2)\delta}$ are defined by

$$\langle \tilde{\delta}_{cl}(\boldsymbol{k}_1)\tilde{\delta}_m(\boldsymbol{k}_2)\tilde{\delta}_m(\boldsymbol{k}_3)\rangle \;=\; (2\pi)^3 \delta_D^{(3)}(\boldsymbol{k}_{123}) B_{c\delta\delta}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3), \tag{71}$$

$$\langle \tilde{\delta}_{cl}(\boldsymbol{k}_1|\nu_{\rm thre,1})\tilde{\delta}_{cl}(\boldsymbol{k}_2|\nu_{\rm thre,2})\tilde{\delta}_m(\boldsymbol{k}_3)\rangle \;=\; (2\pi)^3 \delta_D^{(3)}(\boldsymbol{k}_{123}) B_{c(1)c(2)\delta}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3). \tag{72}$$

At $\ell \gtrsim 500$, the main contributor to the covariances is the one-halo term of the above bi-spectra. Similarly to the case of tri-spectra, the corresponding terms are expressed as

$$B_{c\delta\delta}^{1h}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3; z) \;=\; \int \mathrm{d}M\, \frac{\mathrm{d}n}{\mathrm{d}M}(z, M) \left(\frac{M}{\bar{\rho}_m(z)}\right)^2 S(z, M|\nu_{\rm thre})\tilde{u}_m(\boldsymbol{k}_2|z, M)\tilde{u}_m(\boldsymbol{k}_3|z, M), \tag{73}$$

$$B_{c(1)c(2)\delta}^{1h}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3; z) \;=\; \int \mathrm{d}M\, \frac{\mathrm{d}n}{\mathrm{d}M}(z, M) \left(\frac{M}{\bar{\rho}_m(z)}\right) S(z, M|\nu_{\rm thre,1})S(z, M|\nu_{\rm thre,2})\tilde{u}_m(\boldsymbol{k}_3|z, M). \tag{74}$$

3 NUMERICAL SIMULATION

In order to study in detail the weak lensing statistics considered in this paper, we use mock weak lensing catalogs generated from a set of full-sky weak gravitational simulations. The N-body simulations reproduce the gravity-driven non-Gaussianities in the underlying matter density field, and realistic mask regions are pasted on our mock catalogues. Thus we can compare our model directly with the simulated catalogues.

3.1 N-body simulation

We first run a number of cosmological N-body simulations to generate three-dimensional matter density fields. We use the parallel Tree-Particle Mesh code `Gadget2` (Springel 2005). We arrange the simulation boxes to cover the past light-cone of a hypothetical observer with an angular extent of $4\pi/8$ steradian. A set of simulations consist of six different boxes and cover one-eighth of the sky. In order to cover the full-sky, we use the same boxes by adopting periodic boundary condition (see Figure 2).

The simulations are run with 1024^3 dark matter particles in six different volumes: the box side length ranges from $450\,h^{-1}{\rm Mpc}$ to $2700\,h^{-1}{\rm Mpc}$ with increments of $450\,h^{-1}{\rm Mpc}$. The largest volume simulations with $2700\,h^{-1}{\rm Mpc}$ on a side enable us to simulate the gravitational lensing effect with source redshift of ~ 1. We generate the initial conditions using a parallel code developed by Nishimichi et al. (2009) and Valageas & Nishimichi (2011), which employs the second-order Lagrangian perturbation theory (e.g. Crocce, Pueblas & Scoccimarro 2006). We set slightly different initial redshift $z_{\rm init}$ as the box size increases. In order to generate the initial conditions, we calculate the linear matter transfer function using `CAMB` (Lewis, Challinor & Lasenby 2000). Our fiducial cosmological model is characterized by the following parameters: matter density $\Omega_{\rm m0} = 0.279$, dark energy density $\Omega_{\Lambda 0} = 0.721$, the density fluctuation amplitude $\sigma_8 = 0.823$, the parameter of the equation of state of dark energy $w_0 = -1$, Hubble parameter $h = 0.700$ and the scalar spectral index $n_s = 0.972$. These parameters are consistent with the WMAP nine-year results (Hinshaw et al. 2013). The parameter of our N-body simulations are summarized in Table 1.
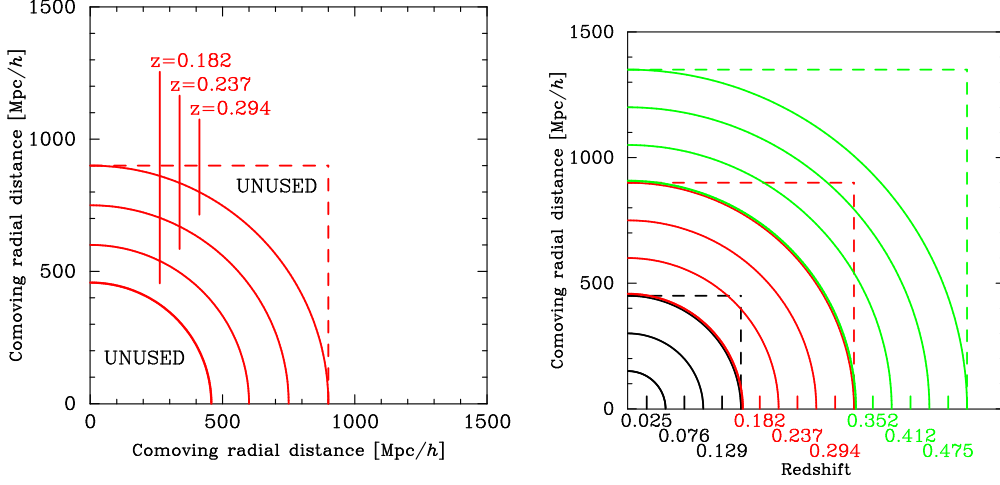
Figure 2. The configuration of spherical shells of projected matter density. The left panel illustrates how we place the outputs of the N-body simulations. There, the dashed line represents the original box size and each of the three shells is taken from the snapshot at different redshift. The right panel shows the nested structure of simulation boxes and the configuration of multiple lens shells. In the right panel, different colors are used to indicate different sets of simulations. The corresponding redshift of each shell is also shown at the bottom of the right panel.

3.2 Ray-tracing simulation

We briefly summarize our ray-tracing simulations with full-sky coverage. The detailed description of our ray-tracing simulations is found in Appendix C. In our ray-tracing simulation, the light ray path and magnification matrix are calculated by the standard multiple lens-plane algorithm. The multiple lens-plane algorithm on a spherical geometry requires contiguous spherical shells of projected matter density. We thus utilize N-body simulations in Section 3.1 to generate a number of thin shells with width of $150\,h^{-1}\mathrm{Mpc}$ and produce three shells from a single simulation box as shown in the left panel of Figure 2. In order to extract a target shell region on a lightcone, we choose an appropriate snapshot at the redshift that correspond to the comoving distance to the shell from a hypothetical observer point (origin). A set of projected density shells are then configured using the nested structure of the simulation boxes. The right panel of Figure 2 shows the configuration of shells used in our ray-tracing simulations (only for $z \lesssim 0.5$).

In this paper, we use `HEALPix` libraries for pixelization of a sphere (Górski et al. 2005). The angular resolution parameter *nside* is set to be 4096. The corresponding angular resolution is $\simeq 0.86$ arcmin. For each shell, we calculate the projected mass density map from N-body particles using Nearest Grid Point method. Then, we perform the spherical harmonic transformation of the density shells to calculate the gravitational lensing potential on each shell via the Poisson equation (see, Eq. (C4)). The obtained gravitational potential and its derivatives are used in the standard multiple lens-plane algorithm. In our simulation, we follow light ray trajectories from $z = 0$ to $z = 1$. The position of a ray and the magnification matrix A_{ij} at each shell are updated by the recurrence relation as in e.g., Hilbert et al. (2009). The initial position of each ray is set to be the center of a pixel on `HEALPix` map. As a ray propagate with deflection, the position at each shell deviates from the position of the `HEALPix` map in general. We evaluate the potential at each shell by using an inverse-distant weighted interpolation of the potential field. The formalism of the multiple lens-plane method on a sphere is found in e.g., Teyssier et al. (2009); Becker (2013).

We obtain a total of ten all-sky convergence maps from ten sets of realization of N-body simulations. Note that we use different initial random seeds to generate N-body simulation data in order to avoid finding similar structures along a line of sight through ray-tracing. Figure 3 shows an example of our simulated convergence map on a full-sky.

3.3 Shape noise and sky mask

We generate realistic mock catalogues by adding the effect of shape noise and sky mask to the obtained convergence maps. The intrinsic ellipticity of source galaxies is the main contaminant of cosmic shear measurement. We model the shape noise in a convergence map by assuming a two-dimensional Gaussian field as follows:

$$\langle \kappa_{\mathrm{N}}(\boldsymbol{\theta}_i)\kappa_{\mathrm{N}}(\boldsymbol{\theta}_j)\rangle = \frac{\sigma_\gamma^2}{2n_{\mathrm{gal}}\Omega_{\mathrm{pix}}}\delta_{ij}, \tag{75}$$

where $\sigma_\gamma = 0.4$, $n_{\mathrm{gal}} = 10\,\mathrm{arcmin}^{-2}$, and $\Omega_{\mathrm{pix}} = 4\pi/12/4096^2 = 6.24 \times 10^{-8}$ str.
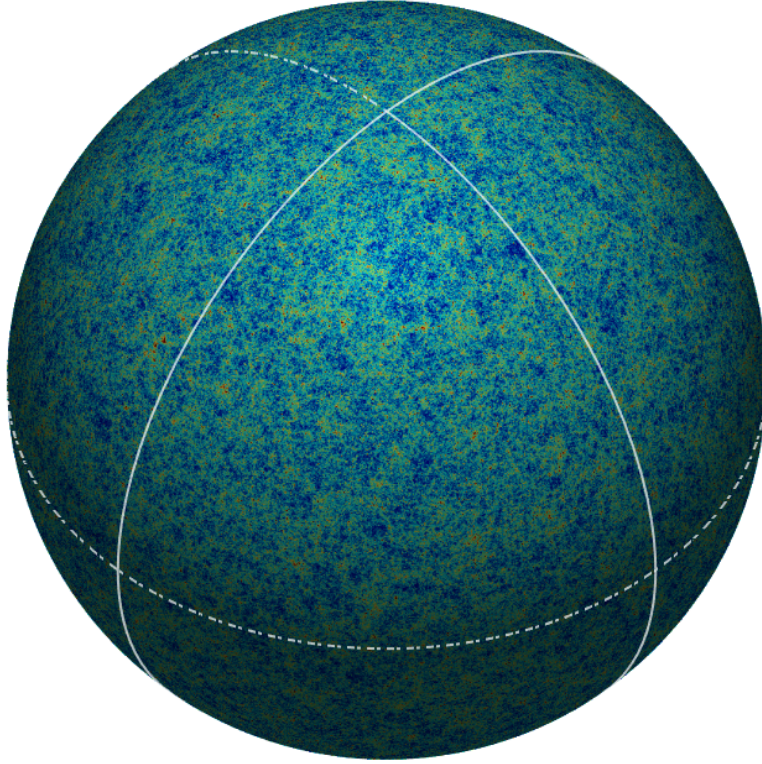
Figure 3. One of our simulated full-sky convergence maps. We perform the Gaussian smoothing with smoothing length of 2.12 arcmin. The shape noise is not included in this figure. The red regions represent high convergence, while the blue regions correspond to negative convergence.
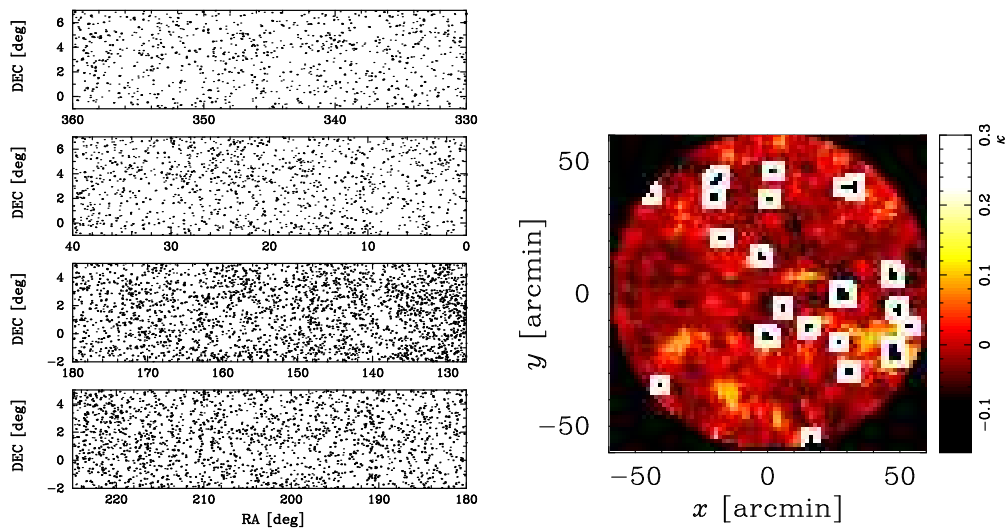

Figure 4. The configuration of the masked regions. In the left panel, black regions show the masks on the bright stars with the R-band magnitude of $M_R < 10$. The right panel represents the typical configuration of the masks in a circular region with radius of 60 arcmin. Black isolated regions show masked regions due to the bright stars and white regions correspond to additional masked region. These white regions remove some pixels affected by the convolution of survey masks (black regions) with the smoothing filter.

It is important to use a realistic sky mask to study statistics for upcoming lensing surveys. We here consider masks owing to bright stars. Among the planned HSC survey regions, we select two continuous regions with sky coverage of ~ 565 and ~ 680 squared degrees. J2000 coordinate of these regions are given by $22^{\rm h}00^{\rm m} < {\rm RA} < 2^{\rm h}40^{\rm m}, -1° < {\rm DEC} < +7°$ and $8^{\rm h}30^{\rm m} < {\rm RA} < 15^{\rm h}00^{\rm m}, -2° < {\rm DEC} < +5°$. In the following, we consider the bright stars with the R-band magnitude of $M_R < 17$. We then select 1,149,871 stars located in these two regions from USNO-A2.0 catalog[4]. In this paper, we assume the following relation between the R-band magnitude M_R and the effective radius of a halo of bright star $r_{\rm star}$:

$$r_{\rm star}[{\rm arcsec}] = \begin{cases} 180 & (M_R < 9), \\ 0.2 \times 10^{[13.75/M_B^{0.7}]} & (9 \le M_R \le 17). \end{cases} \tag{76}$$

Then, we paste a circular mask with $r_{\rm star}$ around the pixel located at each star using Eq. (76). If $r_{\rm star}$ is less than the angular size of our all-sky map (~ 1 arcmin), we simply mask the pixel located in the star. After the above procedure, we further remove the "isolated" pixels whose surrounding pixels are all labeled as masked pixels. The final mask configuration generated in this way is shown in Figure 4. The left panel displays the masked region over the $565 + 680$ square degrees. The black dots shows the masked pixels on very bright stars ($M_R < 10$) and smaller masked regions are distributed in the white region homogeneously (but not shown in the figure). The mask covers over $\simeq 280$ squared degrees in total. The right panel shows an example of our masked sky simulation in a circular region with a radius of $1°$.

From a single full-sky simulation, we make 20 mock HSC convergence maps with the masks by choosing the desired sky coverage ($565 + 680 = 1245$ squared degrees). We allow to have small overlap regions between the 20 masked maps. The overlap regions are located near the edge of the HSC sky coverage, and thus we expect this minor compromise does not affect the final results significantly. Finally, using 10 independent full-sky maps, we obtain a total of $20 \times 10 = 200$ realizations of mock HSC lensing catalogs.

For a smoothed convergence map, we apply the different masking. As shown in Eq. (76), stars with $M_R > 12$ have smaller mask radii than the pixel size in our simulation. Therefore, in principle, we can extract some information from the pixels where stars with $M_R > 12$ are located. Because such faint stars would not affect the smoothed convergence map, we mask only the bright stars with the R-band magnitude of $M_R < 12$. There are also ill-defined pixels that are compromised by the convolution with the smoothing filter. We remove such ill-defined pixels within 5 arcmin from the boundary of the mask regions. Consequently, masked regions on the smoothed convergence $\mathcal{K}$ maps differ from the original masked regions. Finally, we have a total of $\simeq 412$ squared degrees as unmasked regions. The corresponding sky fraction is ~ 0.01.

4   RESULT

We present the weak lensing statistics measured from the full-sky simulations and also those from our 200 mock HSC catalogues. In the following, we define the threshold of lensing peak as $\mathcal{K}/\sigma_{\rm noise,0}$, where $\sigma_{\rm noise,0}$ is given by Eq. (10). Note that we use this definition also in the case without noise.

4.1   Ensemble average of statistics

All-sky

We first show the ensemble average statistics over the ten full-sky simulations. They can be regarded as the expected values from an idealized full-sky observation. Figure 5 summarizes the measurements of $P_{\kappa\kappa}, P_{\rm pp}, P_{\rm p\kappa}$ and $N_{\rm peak}$. We selected the lensing peaks with threshold of $\nu_{\rm thre} = 3$ (or $\mathcal{K} \simeq 0.05$) in both the maps with and without shape noise. To calculate the correlation in harmonic space, we correct the pixelisation effect with the pixel window function of `HEALPix`. Overall, our model is in good agreement with the measurement from the full-sky simulations. In the case without shape noise, we calculate $P_{\rm pp}$, $P_{\rm p\kappa}$ and $N_{\rm peak}$ assuming the one-to-one correspondence between lensing peaks and dark matter halos, i.e.,

$$\text{Prob}(\mathcal{K}_{\rm peak,obs} | \mathcal{K}_{\rm peak,h}(z, M)) = \delta_D(\mathcal{K}_{\rm peak,obs} - \mathcal{K}_{\rm peak,h}(z, M)). \tag{77}$$

The results of $P_{\rm pp}$ and $P_{\rm p\kappa}$ from the maps with and without noise show appreciable differences at large angular scales ($\ell \lesssim 100$). This can be explained by the modulation of peak height due to the shape noise. When we select the lensing peaks with a given $\nu_{\rm thre}$, we effectively include less massive dark matter haloes as well in the noisy $\mathcal{K}$ maps. Then both $P_{\rm pp}$ and $P_{\rm p\kappa}$ are reduced at large angular scales in the noisy maps because less massive halos have weaker clustering. Also, the number count of peaks in the noisy convergence maps is described well by our model as shown in Section 2.3. In the maps with shape noise, we detect more lensing peaks for a given threshold, for example, by a factor of ~ 3 for $\nu_{\rm thre} = 3$.
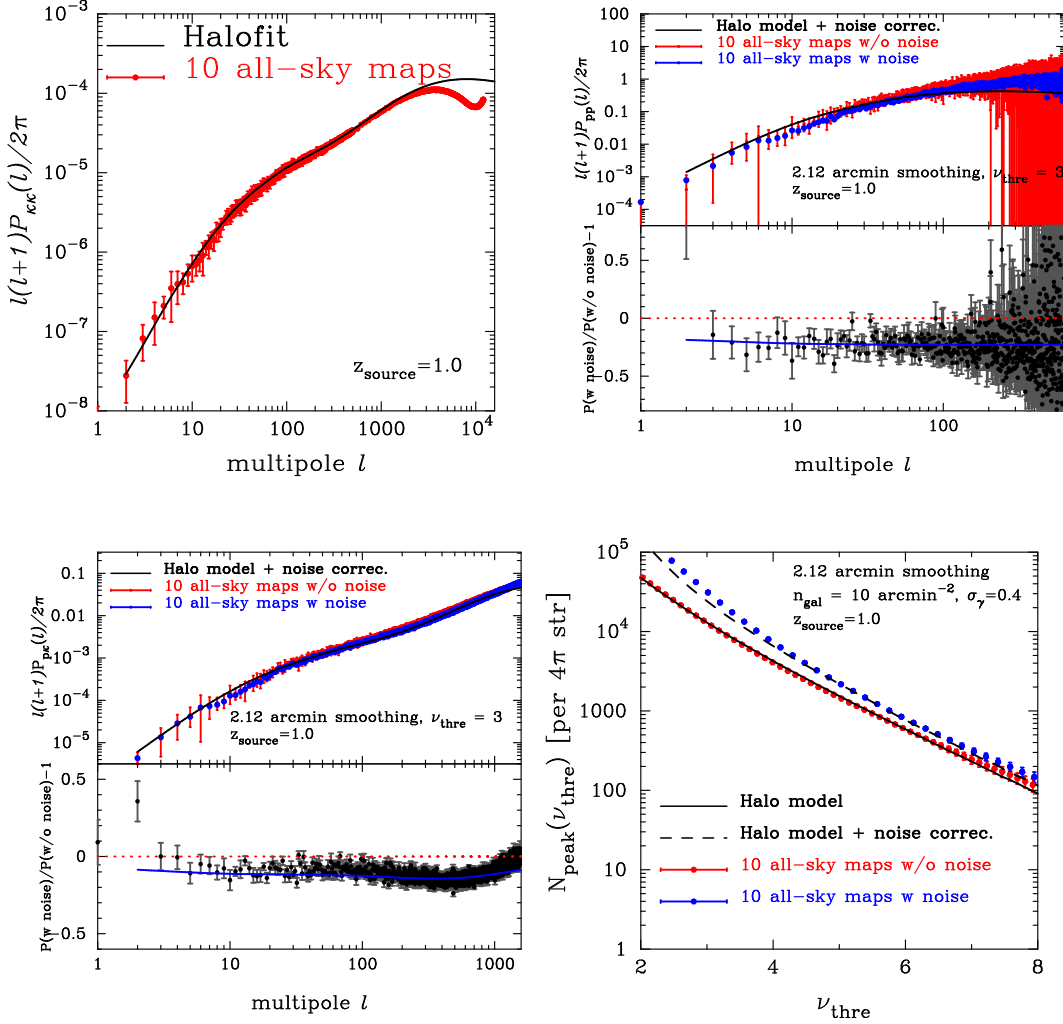
Figure 5. We compare the statistics measured from the ten all-sky simulations and our model prediction. In each panel, the black line is our halo model prediction. The red points with error bar represent the measured signal from the 'clean' convergence maps without noise. The blue points show the result with noise. In the bottom portion in top right and the bottom left panel, we show the ratio of the two. The error bars indicate the standard deviation over ten realization. Note that the threshold of lensing peaks is defined by $\mathcal{K}/\sigma_{\mathrm{noise},0}$ regardless of the presence or absence of noise.

Masked sky

In the presence of masks, we need the correction for the mode-coupling effects in measurements of power spectrum. We adopt the pseudo-spectrum estimator for this purpose (Hansen & Górski 2003; Efstathiou 2004; Brown, Castro & Taylor 2005; Hikage et al. 2011). The basics of the method is summarized in Appendix D. In measurements of the binned two spectra $P_{\kappa\kappa}$ and $P_{\mathrm{p}\kappa}$, we follow the similar method shown in Brown, Castro & Taylor (2005); Hikage et al. (2011). Let us consider the b-th binned power spectrum in the multipole range of $\ell_{\min}^b < \ell < \ell_{\min}^{b+1}$. In this case, we can calculate the band-powers (P_b) as

$$P_b = \sum_{b'} \mathcal{M}_{bb'}^{-1} \sum_{\ell} \mathcal{B}_{b'\ell} \tilde{P}(\ell), \tag{78}$$

where $\tilde{P}(\ell)$ is the measured power spectrum on a masked sky and $\mathcal{B}_{b\ell}$ is the binning operator, which is defined by

$$\mathcal{B}_{b\ell} = \begin{cases} \ell(\ell+1)/(2\pi)/(\ell_{\min}^{b+1} - \ell_{\min}^b) & (\ell_{\min}^b < \ell < \ell_{\min}^{b+1}), \\ 0 & (\text{otherwise}). \end{cases} \tag{79}$$

Here, we define the binned coupling matrix $\mathcal{M}_{bb'}$ as

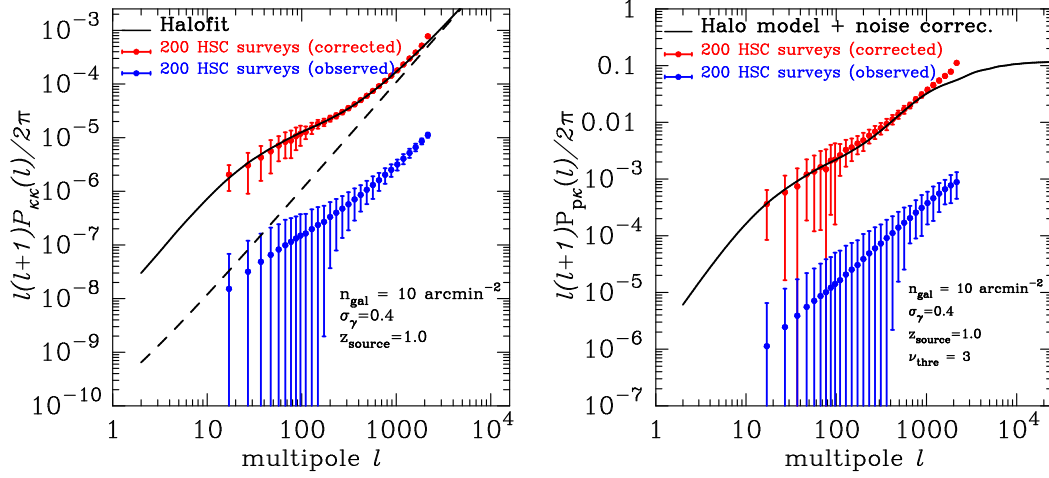$$\mathcal{M}_{bb'} = \sum_{\ell} \mathcal{B}_{b\ell} M_{\ell\ell'} Q_{\ell'b'}, \tag{80}$$

Figure 6. The auto- and cross-power spectra obtained from 200 mock HSC catalogues. The black line is the result of our halo model. In the left panel, the dashed line shows the shape noise contribution to convergence power spectrum.

where the definition of $M_{\ell\ell'}$ is found in Appendix D and $Q_{b\ell}$ is given by

$$Q_{b\ell} = \begin{cases} 2\pi/[\ell(\ell+1)] & (\ell^b_{\min} < \ell < \ell^{b+1}_{\min}), \\ 0 & (\text{otherwise}). \end{cases} \tag{81}$$

The number of bins is set to be 30. We perform the binning in linear spacing for the first 10 bins as $\ell^b_{\min} = 2 + 10b\,(1 \le b \le 10)$, while the remaining 20 bins have logarithmically equal spacing up to $\ell_{\max} = 2000$. The measured and corrected binned power spectra are plotted in Figure 6. We use 200 masked sky simulations to calculate the average of the binned $P_{\kappa\kappa}$ and $P_{p\kappa}$.

For cross-correlation, we find $\sim 100 - 200$ lensing peaks with $\nu_{\text{thre}} = 3$ on each simulated HSC map. In Figure 6, the blue points with error bars shows the measured spectrum, and the red points are the corrected spectrum with the pseudo-spectrum method. Clearly, we can recover the underlying power spectrum with the correction of the mode-coupling due to masks. The difference in amplitude can be explained approximately by the effective fraction of sky coverage ($\sim 0.01 - 0.02$)

4.2   Covariance

We use 200 masked sky simulations to calculate covariances between the weak lensing statistics of interest. First, we calculate the convergence power spectrum covariance. In the flat sky approximation and without masks, the covariance of the binned power spectrum is expressed as (e.g., Cooray & Hu 2001)

$$\text{Cov}[P_{\kappa\kappa}(\ell_i), P_{\kappa\kappa}(\ell_j)] = \frac{\delta_{ij}}{2\ell_i \Delta\ell f_{\text{sky}}} 2P^2_{\kappa\kappa}(\ell_i) + \frac{1}{4\pi f_{\text{sky}}} \int_{\ell_i} \frac{\mathrm{d}^2\boldsymbol{\ell}}{A_{s,i}} \int_{\ell_j} \frac{\mathrm{d}^2\boldsymbol{\ell'}}{A_{s,j}} T_{\kappa\kappa\kappa\kappa}(\boldsymbol{\ell}, -\boldsymbol{\ell}, \boldsymbol{\ell'}, -\boldsymbol{\ell'}), \tag{82}$$

where $P_{\kappa\kappa}(\ell_i)$ represents the i-th binned power spectrum, f_{sky} is the fraction of sky covered by the observation, $\Delta\ell$ is the bin width in ℓ space, and $A_{s,i}$ is the area of the two-dimensional shell around the i-th bin ℓ_i in Fourier space. Here, $T_{\kappa\kappa\kappa\kappa}$ is the tri-spectrum of convergence, of which definition and modeling are found in Section 2.4. In practice, we simplify the second term in Eq. (82) as

$$\int_{\ell_i} \frac{\mathrm{d}^2\boldsymbol{\ell}}{A_{s,i}} \int_{\ell_j} \frac{\mathrm{d}^2\boldsymbol{\ell'}}{A_{s,j}} T_{\kappa\kappa\kappa\kappa}(\boldsymbol{\ell}, -\boldsymbol{\ell}, \boldsymbol{\ell'}, -\boldsymbol{\ell'}) \rightarrow T_{\kappa\kappa\kappa\kappa}(\ell_i, \ell_i, \ell_j, \ell_j). \tag{83}$$

In the presence of masked regions, the estimator of the binned power spectrum is given by more complex expressions given in Eq. (78). The mode-coupling due to masked regions induces the intricate correlation between different Fourier modes in the covariance of $P_{\kappa\kappa}$ (the exact expression is found in Appendix D). Nevertheless, in the case where the value of masked pixels is set to be either 0 or 1, we can use the simplified formula of covariances as shown in Efstathiou (2004). When we work with harmonic space, the approximated formula can be expressed as

$$\text{Cov}[P_{\kappa\kappa}(\ell_1), P_{\kappa\kappa}(\ell_2)]_{\text{mask}} \simeq \sum_{\ell'_1 \ell'_2} \frac{1}{2\ell'_2 + 1} (M^{\kappa\kappa})^{-1}_{\ell_1 \ell'_1} (M^{\kappa\kappa})^{-1}_{\ell_2 \ell'_2} M^{\kappa\kappa}_{\ell'_1 \ell'_2} \left[2P_{\kappa\kappa}(\ell'_1) P_{\kappa\kappa}(\ell'_2) + \frac{1}{4\pi} \sqrt{(2\ell'_1 + 1)(2\ell'_2 + 2)} T_{\kappa\kappa\kappa\kappa}(\ell'_1, \ell'_1, \ell'_2, \ell'_2) \right], \tag{84}$$

where $M^{\kappa\kappa}_{\ell\ell'}$ represents the mode-coupling matrix for convergence. In Eq. (84), we evaluate the term $T_{\kappa\kappa\kappa\kappa}$ by the tri-spectrum in harmonic space following Hu (2001) because we define $T_{\kappa\kappa\kappa\kappa}$ in the flat sky approximation.
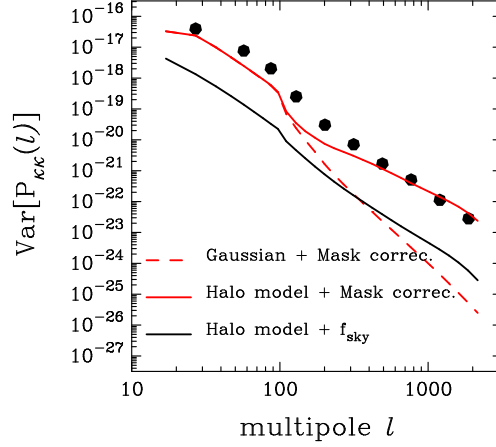
Figure 7. The variance of the psudo-spectrum estimator of $P_{\kappa\kappa}$ measured from 200 masked sky simulations. The black points show our measurement. The three lines represent the different contribution to variance. The black line is the simplest halo model with the sky fraction of $f_{\rm sky} = 0.023$, while the red lines show the halo model covariance with the correction of the effect of masked regions. The red dashed line corresponds to the Gaussian model including the effect of masked region. For the red solid line, we take into account both non-Gaussianities caused by gravity and masked regions.

We can then examine the validity of Eq. (82) by using the measured covariance of $P_{\kappa\kappa}$ over 200 masked sky simulations. We use the same binning as in Section 4.1 but reduce the number of bins to 10 by taking average of the binned powers over nearest ℓ bins. Figure 7 shows the measured variance of the pseudo-spectrum estimator of $P_{\kappa\kappa}$. The black point shows the result obtained from the 200 simulations and the solid line represents our model of covariance in Eq. (82) with the appropriate value of $f_{\rm sky}$ for our simulations. Interestingly, although the simple model of Eq. (82) is expected to account for non-Gaussianities caused by gravity, it underestimates the actual covariance by a factor of ~ 10. The corrected covariance components are shown in Figure 7. The first term and second term in Eq. (84) are plotted as red dashed and red solid line, respectively. The overall amplitude of the variance can be explained by the mask correction, while the contribution from tri-spectrum dominates at $\ell \gtrsim 200$. Clearly, it is problematic to adopt the commonly used estimate of covariance given by in Eq. (82).

To perform combined analysis of $P_{\kappa\kappa}, P_{\rm p\kappa}$ and $N_{\rm peak}$, we need cross covariance between the statistics. Unfortunately, it is difficult to model the impact of masked region on the cross covariances in a similar manner to Eq. (84). Here, we simply compare the measured cross covariances with the prediction of our halo model. In the model, the covariance can be derived similarly to Eq. (82). For example,

$$\mathrm{Cov}[P_{\rm p\kappa}(\ell_i), P_{\rm p\kappa}(\ell_j)] \quad = \quad \frac{\delta_{ij}}{2\ell_i \Delta\ell f_{\rm sky}} \left(P_{\rm pp}(\ell_i)P_{\kappa\kappa}(\ell_i) + P_{\rm p\kappa}(\ell_i)P_{\rm p\kappa}(\ell_i)\right) + \frac{1}{4\pi f_{\rm sky}} \int_{\ell_i} \frac{\mathrm{d}^2\boldsymbol{\ell}}{A_{s,i}} \int_{\ell_j} \frac{\mathrm{d}^2\boldsymbol{\ell}'}{A_{s,j}} T_{\rm p\kappa p\kappa}(\boldsymbol{\ell}, -\boldsymbol{\ell}, \boldsymbol{\ell}', -\boldsymbol{\ell}'), \quad (85)$$

$$\mathrm{Cov}[P_{\rm p\kappa}(\ell_i), P_{\kappa\kappa}(\ell_j)] \quad = \quad \frac{\delta_{ij}}{2\ell_i \Delta\ell f_{\rm sky}} 2P_{\rm p\kappa}(\ell_i)P_{\kappa\kappa}(\ell_i) + \frac{1}{4\pi f_{\rm sky}} \int_{\ell_i} \frac{\mathrm{d}^2\boldsymbol{\ell}}{A_{s,i}} \int_{\ell_j} \frac{\mathrm{d}^2\boldsymbol{\ell}'}{A_{s,j}} T_{\rm p\kappa\kappa\kappa}(\boldsymbol{\ell}, -\boldsymbol{\ell}, \boldsymbol{\ell}', -\boldsymbol{\ell}'), \quad (86)$$

where we have defined $T_{\rm p\kappa p\kappa}$, $T_{\rm p\kappa\kappa\kappa}$, and the other covariances (i.e. $\mathrm{Cov}[N_{\rm peak}, N_{\rm peak}]$, $\mathrm{Cov}[N_{\rm peak}, P_{\kappa\kappa}]$, and $\mathrm{Cov}[N_{\rm peak}, P_{\rm p\kappa}]$) in Section 2.4.

Figure 8 shows the measured peak-power cross-covariance. In constrast to the case of $P_{\kappa\kappa}$, the covariance including $P_{\rm p\kappa}$ and $N_{\rm peak}$ is less affected by masks: the simple model that accounts for $f_{\rm sky}$ yields a reasonable result with respect to that of the simulations. We find that the difference is by a factor of ~ 3 at most.

5 CONCLUSION AND DISCUSSION

We have performed all-sky lensing simulations to generate a large set of realistic weak lensing mass maps with complex masked regions by incorporating the actual position of bright stars. We have used the set of the mock samples to study the statistical properties of weak lensing convergence and convergence peaks in detail. Full nonlinear covariances between the statistics have been also calculated from 200 realization of masked maps. We have also developed an analytic halo model that provides reasonably accurate prediction for the statistics.

When adopting a Gaussian smoothing with the full width at half maximum of 5 arcmin, we can associate weak lensing convergence peaks with dark matter halos with mass of $\sim 10^{14} h^{-1} M_{\odot}$ at $z \sim 0.1 - 0.2$. We can also estimate the modulation
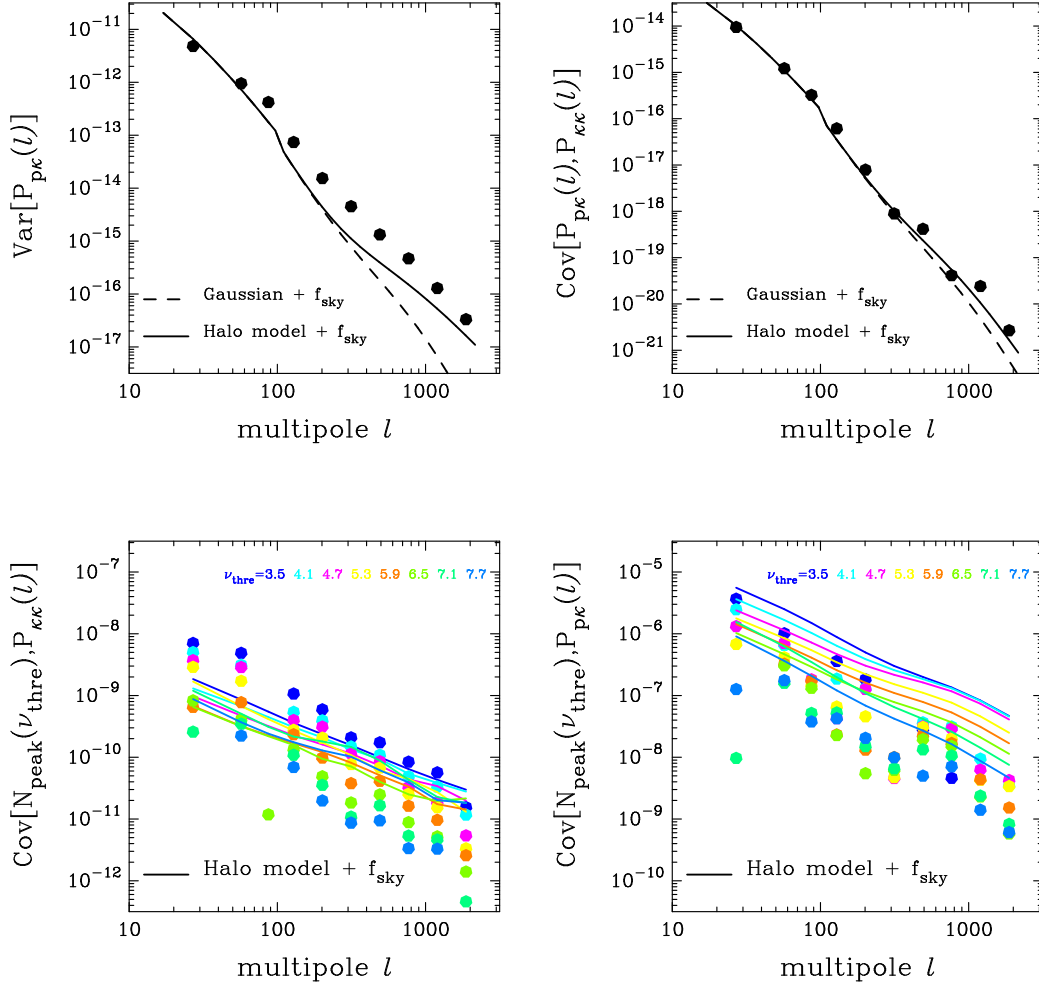
Figure 8. The cross covariance of weak lensing statistics calculated from the 200 simulations. In each panel, the black or colored points show the measured covariance, and the solid line shows the corresponding halo model prediction. We define the peaks with ν_{thre} when measuring $P_{\mathrm{p}\kappa}$. In the bottom two panels, the color indicate the results for different threshold ν_{thre} in N_{peak}. The threshold value is varied in the range $\nu_{\mathrm{thre}} = 3.5 - 7.7$.

of peak height due to shape noise by using a model based on Gaussian peak statistics. Thus, the abundance of peaks, the angular correlation function, and the cross-correlation of peaks and cosmic shear are all obtained accurately by our halo model approach with the shape noise correction. Furthermore, the halo model can also take the mask effect into account and indeed produces accurate ensemble average of statistics and their covariances. The impact of masked regions on the covariance can be described by the following two effects: (i) reduction of sky coverage and (ii) the mode-coupling effect between different Fourier modes. We find that the former affects the overall amplitude of cross-covariance between statistics, while the latter is important for the covariance of cosmic shear power spectrum $P_{\kappa\kappa}$. For the masked regions adopted here, ignoring the mode-coupling effect would induce underestimation of the covariance of $P_{\kappa\kappa}$ by a factor of $\sim 10(!)$.

The number density of source galaxies is an important factor in the statistical analysis of weak gravitational lensing. As one may expect, a large number density of sources is desired to find clusters with high accuracy and perform cosmological analysis with selected clusters. In comparison with the case of $n_{\mathrm{gal}} = 10\,\mathrm{arcmin}^{-2}$, we have confirmed that the signal-to-noise ratio increase by a factor of ~ 1.5 in combined analysis with $P_{\kappa\kappa}$, $P_{\mathrm{p}\kappa}$ and N_{peak} even if we ignore the shape noise contaminant (i.e., $n_{\mathrm{gal}} \to \infty$). This suggests that imaging over a wide area is suitable for cosmological analysis with the lensing statistics even if the number density of sources is not significantly increased in such surveys.

We have also examined the validity of our model for two additional cases: $n_{\mathrm{gal}} = 5$ and $30\,\mathrm{arcmin}^{-2}$. When the smoothing scale, the rms of intrinsic ellipticity, and the source redshift are all fixed, the halo model prediction is in good agreement with

the result of our full-sky simulations in the case of $n_{gal} = 30\, arcmin^{-2}$, but the agreement is worse with $n_{gal} = 5\, arcmin^{-2}$ [5]. Therefore, when we consider the typical value of the rms of intrinsic ellipticity and the source redshift, our model is expected to be accurate when $n_{gal} \gtrsim 10\, arcmin^{-2}$ with ~ 2 arcmin Gaussian smoothing.

Our model has been successfully applied the statistics of a smoothed convergence map with shape noise. In principle, one can find an optimal filter function so that the number of detected clusters is increased (e.g. Hennawi & Spergel 2005; Maturi et al. 2005). Our model is based on the assumption that shape noise in a smoothed lensing map has Gaussian properties. The assumption is valid when the ellipticities of the source galaxies are uncorrelated and when there are a sufficient number of source galaxies per pixel (i.e. the central limit theorem). We expect that, while our model can be applied to the general form of filter function, it may not be appropriate when there are only few source galaxies and/or when there is no one-to-one correspondence between peaks and halos.

Clusters of galaxies are important targets in future cosmological surveys. Selecting galaxy clusters based on cosmic shear measurement does not rely on mass estimate nor on calibration with additional information such as X-ray brightness. Nevertheless, there remain some systematic effects worth mentioning here.

Gravitational lensing causes not only distortion of source images but also magnification. Using magnified galaxies in a flux- and size-limited survey would potentially cause systematic effect(s) in the weak lensing statistics. The magnification effects on lensing peak statistics have been already studied in e.g., Schmidt & Rozo (2011). Recent numerical study by Liu et al. (2014a) suggests that the magnification effect causes non-negligible bias in parameter estimation in the case of LSST. In order to examine the magnification effect further, it is essential to run high angular resolution simulations. This is because the mean number of source galaxies on each pixel should be less than unity to make the one-to-one correspondence between a pixel and a (magnified) source galaxy. In the case of $n_{gal} = 10\, arcmin^{-2}$, we should set the pixel size to be $1/\sqrt{n_{gal}} \sim 0.3\, arcmin$. We will perform such simulations to study the magnification effect in wide-field surveys in detail.

Another important issues are uncertainties and systematic bias associated with baryonic effects. Previous studies (e.g., Semboloni et al. 2011; Semboloni, Hoekstra & Schaye 2013; Zentner et al. 2013) explored the impact of the baryonic component to two-point statistics of cosmic shear and consequently to cosmological parameter estimation. The baryonic effect is likely important in weak lensing peak statistics. Indeed, Yang et al. (2013) show appreciable baryonic effects on peak statistics using a simple model applied to dark-matter-only simulations, whereas the baryonic effect on higher order convergence statistics have been studied with numerical simulations (Osato, Shirasaki & Yoshida 2015). Recently, Mohammed et al. (2014) explored halo model approach to include the baryonic effect on cosmic shear statistics.

The statistical properties and the intrinsic correlation of source galaxies and lensing structures are still uncertain but could be critical when making a large lensing mass maps. Among such correlations, source-lens clustering (e.g., Hamana et al. 2002) and the intrinsic alignment (e.g., Hirata & Seljak 2004) are likely to compromise cosmological parameter estimation. A promising approach in theoretical studies would be associating the source positions with their host dark matter halos on the light cone. This is along the line of our ongoing study using a large set of cosmological simulations.

Weak gravitational lensing is a promising tool to probe the dark matter distribution in the universe. Statistical analysis of a reconstructed mass map can be performed to extract precise cosmological information. The peak statistics considered in the present paper contain the information related to massive objects such as clusters of galaxies and thus have a great potential to probe cosmology and constrain the model of structure formation simultaneously. Ongoing/upcoming imaging surveys such as HSC, DES, and LSST in the near future, will provide the largest dark matter map we have never seen before. We expect our study presented here provides a useful guide to interpret properly the reconstructed mass map and to reveal the nature of the dark components in the universe.

APPENDIX A: HALO-PEAK MATCHING

In this appendix, we examine the correspondence between dark matter halos and the local maximum in lensing convergence map.

---

[5] We expect that the disagreement for small n_{gal} would be caused by the offset between the position of a peak and the center of the corresponding halo. The offset effect would be more important when shape noise increases as shown in Fan, Shan & Liu (2010).
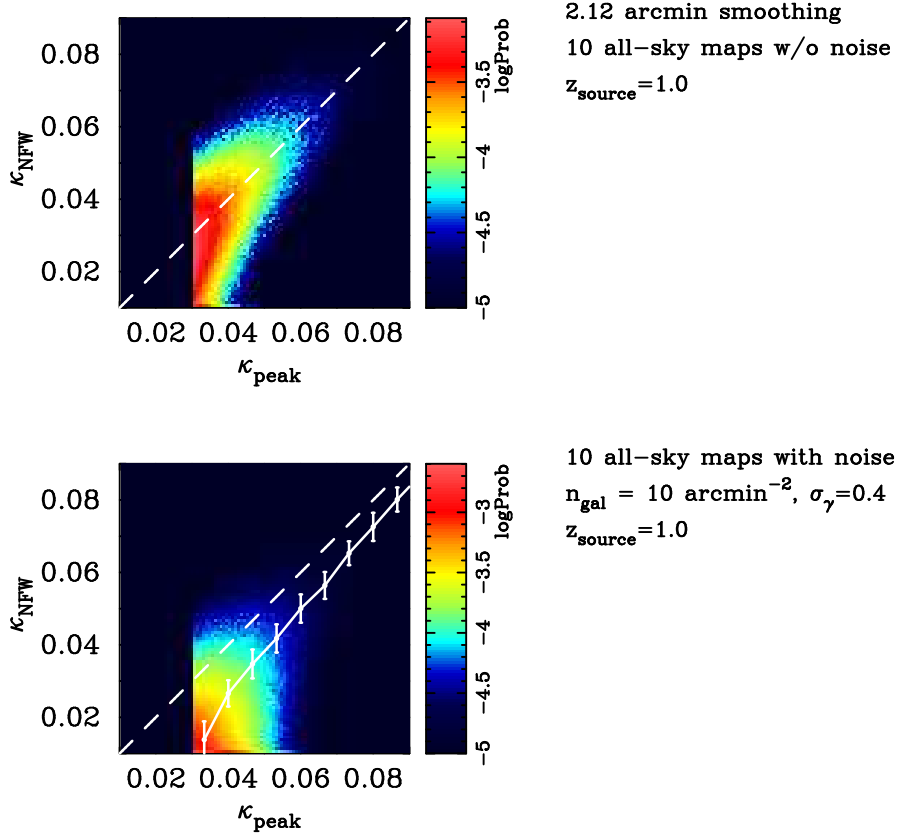
Figure A1. The correspondence between dark matter halos and lensing peaks. The top panel shows the scatter plot of peak height and the expected convergence by the matched halos in absence of noise. The lower panel corresponds to the case with shape noise. In each panel, the horizontal axis represents the peak height and the vertical axis shows the expected convergence of NFW halos.

In order to generate mock halo catalogs, we identify dark matter halos in outputs of our N-body simulation (see, Section 3.1) using the standard friends-of-friends algorithm with a linking parameter of $b = 0.2$ in units of the mean particle separation. We use dark matter halos with mass greater than $10^{13}\,h^{-1}M_\odot$ in the following analysis. This lowest mass corresponds to the mass of 20 particles in the largest simulation. Then, the position of dark matter halo in N-body simulations are arranged in the same way as in the ray-tracing experiments in Section 3.2.

With our ray-tracing simulations and mock halo catalogs, we study the correspondence between halos and the peaks in weak lensing convergence maps. We first identify the local maxima in the smoothed lensing convergence field with source redshift of $z_{\rm source} = 1$. In this appendix, we again adopt the Gaussian smoothing with the full width at half maximum of 5 arcmin. When including the shape noise in convergence maps, we set $\sigma_\gamma = 0.4$ and $n_{\rm gal} = 10\,{\rm arcmin}^{-2}$. For selection of peaks, the threshold of peak height is set to be $\mathcal{K} = 0.03$. This value corresponds to $\sim 3\sigma$ in smoothed convergence maps without noise. For a given position of lensing peak, we search for the matched dark matter halos within a radius of 5 arcmin from the peak position. This search radius is set to be larger than the smoothing scale but still smaller than the angular size of massive halos at $z \sim 0.1 - 0.2$ (also see, Hamana, Takada & Yoshida 2004). When we find several halos in search radius, we regard the matched halo as the closest halo from the position of peak. For each matched peak, we estimate the corresponding convergence by using the universal NFW density profile (see Section 2.2 in detail). In the calculation of expected convergence from FoF halos, we simply assume that the FoF mass is equal to the virial mass. In total, we find 632,238 and 1,404,538 pairs of peaks and halos over 10 noise-less maps and noisy maps, respectively.

Figure A1 shows the scatter plot of peak height in $\mathcal{K}$ map and the expected convergence by NFW halos. The horizontal axis corresponds to peak height, while the vertical axis shows the corresponding convergence expected by NFW halos. Thus, the color map in each panel shows the probability of Eq. (17). We present the line of $y = x$ as the dashed line in each panel. In lower panel of this figure, we show the effect of the modulation of peak height as the solid line with error bars. The solid line is derived by $\bar{\mathcal{K}}_{\rm peak,obs}(z, M)$ in Eq. (22) and the error bars reflect the scatter of $\bar{\mathcal{K}}_{\rm peak,obs}(z, M)$. As shown in previous works, we confirm the good correspondence between the matched dark matter halos and lensing peaks in the noise-less maps.
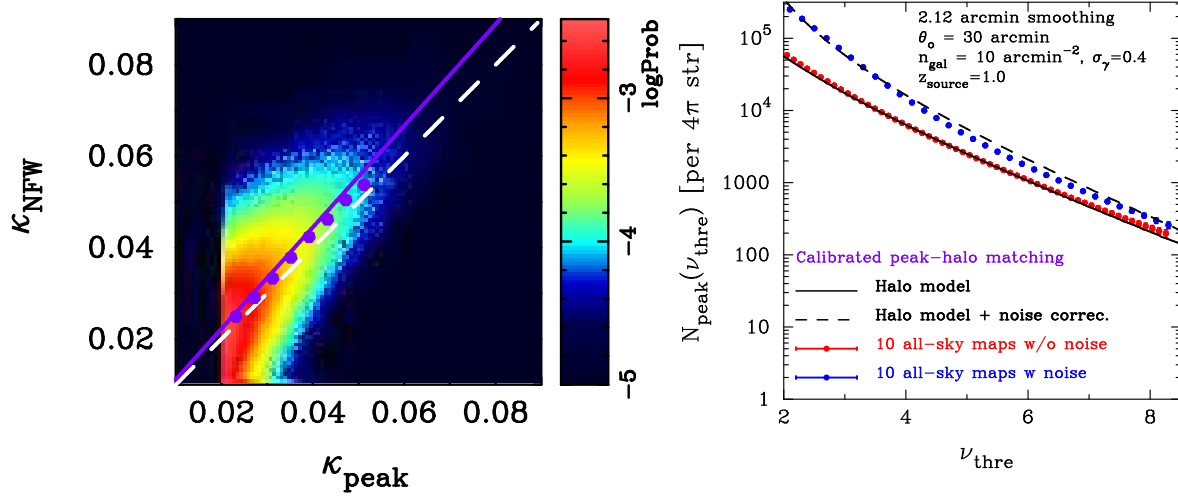
Figure B1. The property of lensing peaks on smoothed maps by the compensated Gaussian filter. The left panel shows the scatter plot of peak height and the expected convergence by the matched halos in absence of noise. In the left panel, the dashed line shows one-to-one correspondence. The purple points represent the mean relation of the expected convergence and measured peak height and the best-fit linear relation is expressed as the purple line. The right panel shows the comparison the peak abundance measured from ten full-sky simulations and our model prediction. In the right panel, the black line is our halo model prediction. The red points with error bar represent the measured signal from the 'clean' convergence maps without noise. The blue points show the result with noise. We normalize the measured peak height by $\sigma_{\mathrm{noise},0}$ in both cases with and without noise. The error bars indicate the standard deviation over ten realization. In the calculation of peak count, we correct the biased relation of κ_{NFW} and κ_{peak} as shown in the purple line in the left panel. The details of correction are found in the text.

Also, our model as shown in Eq. (22) can explain the average relation between peaks and dark matter halos even in the case with noise.

APPENDIX B: THE CASE OF COMPENSATED GAUSSIAN FILTER

We examine another practical filter function to construct smoothed lensing maps. In this appendix, we consider $z_{\mathrm{source}} = 1$, and we set $\sigma_\gamma = 0.4$ and $n_{\mathrm{gal}} = 10\,\mathrm{arcmin}^{-2}$ when including the shape noise in convergence maps.

We first consider the compensated filter U_{c} based on the Gaussian form as

$$U_{\mathrm{c}}(\theta) = \frac{1}{\pi\theta_G^2}\exp\left(-\frac{\theta^2}{\theta_G^2}\right) - \frac{1}{\pi\theta_o^2}\left[1 - \exp\left(-\frac{\theta_o^2}{\theta_G^2}\right)\right], \tag{B1}$$

where θ_o represents the boundary of the filter and we set U_{c} to be zero for $\theta > \theta_o$. We adopt the smoothing scale of $\theta_G = 5/\sqrt{8\ln 2}$ arcmin and $\theta_o = 30$ arcmin. For the compensated filter function of U_{c}, the noise power spectrum on a smoothed lensing map is expressed as (van Waerbeke 2000)

$$P_{\mathcal{N}}(\ell) = \frac{\sigma_\gamma^2}{2n_{\mathrm{gal}}}|\tilde{U}_{\mathrm{c}}(\ell)|^2, \tag{B2}$$

where σ_γ is the rms of the intrinsic ellipticity of sources, n_{gal} represents the number density of source galaxies, and $\tilde{U}_{\mathrm{c}}$ is the Fourier transform of U_{c}. As shown in Eq. (10), the noise variance $\sigma_{\mathrm{noise},0}$ is evaluated by the integral of $P_{\mathcal{N}}(\ell)$ in Fourier space. In the case of $\theta_o = 30$ arcmin, the boundary of the filter changes $\sigma_{\mathrm{noise},0}$ by about 1% compared to the case of the usual Gaussian filter. Thus, we can safely ignore the difference of $\sigma_{\mathrm{noise},0}$ between the compensated and the usual Gaussian filter for our parameter choice.

In order to investigate the effect of the modification of filter function on lensing peak statistics, we first study the correspondence between peaks and halos by the halo-peak matching analysis as shown in Appendix A. When we limit lensing peaks with the height larger than 0.02, we find 1,045,291 matched pairs over ten full-sky maps without shape noise. The left panel in figure B1 shows the scatter plot of the measured peak height and the expected convergence signal for the spherical NFW halo. In the calculation of expected signal, we simply assume that FoF mass of halos is equal to the virial mass and use the model of concentration parameter c_{vir} in Duffy et al. (2010). We expect that the one-to-one correspondence between peaks and halos would still hold. However, we find biased relation between the mean measured height and expected one *even in the absence of noise*. The biased mean relation is shown by the purple points in the left panel in figure B1. It can also be fitted

well by linear relation of $\mathcal{K}_{\mathrm{peak}} = \alpha \mathcal{K}_{\mathrm{NFW}} + \beta$. We find the best-fit value of α is 0.9 while the offset β can be approximated to be zero. A similar relation is also found by Hamana et al. (2012). It might be caused as a consequence of various effects such as the mismatch of FoF mass and virial mass defined by spherical over-density.

Even without knowing the origin of the biased relation, we can still predict the lensing peak statistics with the compensated filter by adopting the biased relation in our halo model. Our approach is simply to replace $\mathcal{K}_{\mathrm{peak},h}(z, M)$ with $\alpha \mathcal{K}_{\mathrm{peak},h}(z, M) + \beta$ for the calculation of Eq. (17). Here, we also assume that $\mathcal{K}_{\mathrm{peak},h}$ can be evaluated by

$$\mathcal{K}_{\mathrm{peak},h}(z, M) = \int \mathrm{d}^2\theta \, U_{\mathrm{c}}(\theta; \theta_G, \theta_o) \kappa_h(\theta|z, M), \tag{B3}$$

where $\kappa_h(\theta|z, M)$ represents the convergence profile of spherical NFW halos with mass of M and the redshift of z. With the above correction, our model can provide a reasonable fit to the measured peak statistics from ten full-sky maps as shown in the right panel in figure B1. The right panel shows the measured peak count for the compensated filter with and without shape noise. In the right panel, the black solid (dashed) line represent our halo model in absence (presence) of noise with the correction for the biased relation of $\mathcal{K}_{\mathrm{peak}}$ and $\mathcal{K}_{\mathrm{NFW}}$. With the above suitable modifications, our model works as long as the one-to-one correspondence of peaks and halos holds. We simply need to calibrate the mean scaling relation of $\mathcal{K}_{\mathrm{peak,obs}}$ and $\mathcal{K}_{\mathrm{peak},h}$ in absence of noise.

APPENDIX C: FULL SKY RAY-TRACING SIMULATION

Here we first summarize basic equations of the multiple-plane gravitational lensing algorithm, and then describe the ray-tracing method through the multiple-plane. For the former we largely follow Das & Bode (2008), and for the latter we adopt one developed by Teyssier et al. (2009).

Throughout this section, we work on the comoving coordinates. Thus ρ, χ, and $r(\chi)$ denote the comoving matter density, the radial comoving distance, and the comoving angular diameter distance, respectively.

C1   Construction of the lensing potentials of multiple-plane

The 3-dimensional light-cone matter distribution is composed of the multiple-layer of shells with a fixed width of $150\,h^{-1}\mathrm{Mpc}$ taken from the nested simulation boxes. The surface matter density field on a sphere for j-th shell is defined by

$$\Delta^j_\Sigma(\boldsymbol{\theta}) = \int_{\mathrm{shell}} \mathrm{d}\chi \, (\rho(\boldsymbol{\theta}, \chi) - \bar\rho) r(\chi)^2, \tag{C1}$$

where $\boldsymbol{\theta}$ denote the angular directions and $\bar\rho$ corresponds the mean matter density, and the integration is over the shell width. We set the lens-planes for each shell at the cone-volume weighted mean distance, $\chi_j = 0.75(\chi^4_{j,\mathrm{max}} - \chi^4_{j,\mathrm{min}})/(\chi^3_{j,\mathrm{max}} - \chi^3_{j,\mathrm{min}})$, where $\chi_{j,\mathrm{max}}$ and $\chi_{j,\mathrm{min}}$ are the farthest and nearest radial distance to a shell, respectively. The convergence field for j-th shell is given by

$$K^j(\boldsymbol{\theta}) = \frac{4\pi G}{c^2} \frac{\Delta^j_\Sigma(\boldsymbol{\theta})}{a_j r(\chi_j)} \tag{C2}$$

where a_j is the scale factor at the lens-plane χ_j. We use the `HEALPix` (Górski et al. 2005) scheme for pixelization of a sphere, and we make the best use of the `HEALPix` library. For each shell, we construct the projected mass density map from N-body particles using Nearest Grid Point method (implemented with `HEALPix` subroutine `vec2pix_ring`). Using the volume (V_{sim}) and total number of N-body particles (N_{part}) of the N-body simulations, the total number of pixels ($npix$), the number of N-body particles within i-th pixel ($n_{\mathrm{pix},i}$) and its mean value ($\bar{n}_{\mathrm{pix}}$), the convergence field is given by

$$K^j(\boldsymbol{\theta}_i) = \frac{3\Omega_m}{2a_j r(\chi_j)} \left(\frac{H_0}{c}\right)^2 \frac{V_{\mathrm{sim}}}{N_{\mathrm{part}}} \frac{npix}{4\pi}(n_{\mathrm{pix},i} - \bar{n}_{\mathrm{pix}}). \tag{C3}$$

Having the convergence field being ready, we expand it in spherical harmonics to have its coefficients, K^j_{lm}, using the `HEALPix` subroutine `map2alm`. Then, the spherical harmonics coefficients for the lensing potential ϕ^j can be obtained via

$$\phi^j_{lm} = \frac{2}{l(l+1)} K^j_{lm} \quad \text{for } l \neq 0, \tag{C4}$$

and $\phi^j_{lm} = 0$ for $l = 0$. This gives us the lensing potential field on a sphere, and its 1st and 2nd derivatives relate to the gravitational lensing deflection field ($\alpha^j_i = -\nabla_{\hat{n}_i}\phi^j$) and the optical tidal matrix ($U^j_{ik} = -\nabla_{\hat{n}_i}\nabla_{\hat{n}_j}\phi^j$), respectively. Note that $\nabla_{\hat{n}_i}$ ($i = 1, 2$) denotes the angular derivative. In an actual computation, we utilize the `HEALPix` subroutine `alm2map_der`.

C2   Light ray propagation

Let us first describe the method to trace the ray trajectory using the multiple-plane algorithm, for which we basically follow one developed by Teyssier et al. (2009). A virtual observer is located at the center of the nested simulation boxes. Rays are traced backward from the observer point with the initial ray directions being set on `HEALPix` pixel centers. Thus the ray positions on the 1st (closest to the observer) is exactly at the `HEALPix` pixel centers. At each lens-plane, the ray directions are deflected according to α_i^j, and the ray positions on the next lens-plane are computed using the method described in Appendix A of Teyssier et al. (2009). Note that ray positions on the j-th ($j > 1$) lens-plane are not exactly at the pixel center due to the lensing deflections, however the lensing fields (α_i^j, U_{ik}^j) are only computed at the pixel centers. In order to evaluate the lensing fields at an arbitrary position ($\boldsymbol{\theta}$), we adopt the inverse distance weighted interpolation from nearest four pixel values;

$$\alpha_i^j(\boldsymbol{\theta}) = \frac{\sum_{k=1}^{4} w_k \alpha_i^j(\boldsymbol{\theta}_k \to \boldsymbol{\theta})}{\sum_{k=1}^{4} w_k}, \tag{C5}$$

where $w_k = 1/|\boldsymbol{\theta} - \boldsymbol{\theta}_k|$, and $\alpha_i^j(\boldsymbol{\theta}_k \to \boldsymbol{\theta})$ is the deflection angle at the pixel center $\boldsymbol{\theta}_k$ but after parallel transporting to the ray position $\boldsymbol{\theta}$ (in order to take into account the change in the local ($\boldsymbol{e}_\theta, \boldsymbol{e}_\phi$) basis). In the actual computation, the parallel transport of the vector and tensor (for U_{ik}^j described below) is implemented by the rotation of them by an angle between two coordinate bases at $\boldsymbol{\theta}_k$ and $\boldsymbol{\theta}$ (see Appendix C of Becker 2013).

Having evaluated the ray positions at a lens-plane, we are able to compute the lensing magnification matrix (A_{ik}^j) using the recurrence relation (Hilbert et al. 2009; Becker 2013);

$$
\begin{aligned}
A_{ik}^{j+1} &= \left(1 - \frac{r(\chi_j)}{r(\chi_{j+1})} \frac{r(\chi_{j+1} - \chi_{j-1})}{r(\chi_j - \chi_{j-1})}\right) A_{ik}^{j-1} + \frac{r(\chi_j)}{r(\chi_{j+1})} \frac{r(\chi_{j+1} - \chi_{j-1})}{r(\chi_j - \chi_{j-1})} A_{ik}^j - \frac{r(\chi_{j+1} - \chi_j)}{r(\chi_{j-1})} U_{im}^j A_{mk}^j, \\
A_{ik}^1 &= \delta_{ik}, \\
A_{ik}^0 &= \delta_{ik},
\end{aligned}
\tag{C6}
$$

for $j \geq 1$, and note that in our notation, the lens-plane closest to the observer is $j = 1$. The optical tidal matrix, U_{ik}^j, in the above relation is evaluated at the ray position on each lens-plane in the same interpolation scheme as eq. (C5), and then again parallel transporting to the unperturbed ray position (i.e., the initial ray direction), because the observed magnification matrix should be evaluated in the local basis of the image position.

We choose the source-plane at an arbitrary redshift z_s (and thus the correspondence radial distance to the source-plane χ_s), and evaluate the source position on the source-plane and the magnification matrix by using the above methods but replacing, e.g., $\chi_{j+1} \to \chi_s$.

C3   Image positions of haloes

The 3-dimensional light-cone distribution of dark matter haloes is generated in the same manner as for the matter distribution. The spatial position of a halo is converted into the angular position $\boldsymbol{\theta}_S^{\mathrm{halo}}$, where the subscript "$s$" means the *source position*. We search for the corresponding *image position* $\boldsymbol{\theta}_I^{\mathrm{halo}}$ in the following manner. First, we search for the nearest ray to the halo source position on the lens-plane of the shell where the halo is located. The displacement vector between the angular positions of the halo and the nearest ray is computed, $\boldsymbol{\Delta\theta_S} = \boldsymbol{\theta}_S^{\mathrm{halo}} - \boldsymbol{\theta}_S^{\mathrm{ray}}$. This vector is parallel transported to the image position of the nearest ray $\boldsymbol{\theta}_I^{\mathrm{ray}}$, and we denotes it by $\boldsymbol{\Delta\theta_I}$ Then the image position of the halo is given by $\boldsymbol{\theta}_I^{\mathrm{halo}} = \boldsymbol{\theta}_I^{\mathrm{ray}} + \boldsymbol{\Delta\theta_I}$. The last step is valid if the difference in the lensing deflection angles between ray-trajectory to the halo and the nearest ray is very small. The statistical properties of differences in the lensing deflection angles between nearby two rays (the, so-called, the lensing excursion angle) were studies in Hamana & Mellier (2001); Hamana et al. (2005). They found that the root-mean-square (rms) value of the lensing excursion angles of rays for $z_s = 1$ with the separation of 1 arcmin is ~ 1 arcsec. This value can be considered as the typical error in $\boldsymbol{\theta}_I^{\mathrm{halo}}$. Considering the fact that the pixel scale of the current ray-tracing simulation is ~ 1 arcmin, we may conclude that the above approximation is reasonably valid. However it should be noticed that for rays gone through a strong lensing region, the excursion angle can be much larger than the rms value, and thus $\boldsymbol{\theta}_I^{\mathrm{halo}}$ may not be very accurate. There is room for improvement on this issue that we leave for future work.

APPENDIX D: STATISTICAL PROPERTY OF PSEUDO-SPECTRUM ESTIMATORS

In this appendix, we summarize the statistical property of pseudo-spectrum estimators. The pseudo-spectrum method is a powerful framework to construct the power spectrum of an underlying random field on limited sky (e.g., Hansen & Górski 2003; Efstathiou 2004; Brown, Castro & Taylor 2005).

Let us consider the two random fields in each direction in the sky: convergence field $\kappa(\Omega)$ and number density field of

lensing peaks $p(\Omega)$. These two fields would commonly be expanded in spherical harmonic as follows:

$$\kappa(\Omega) = \sum_{\ell m} \kappa_{\ell m} \mathcal{Y}_{\ell m}(\Omega), \tag{D1}$$

where $\mathcal{Y}_{\ell m}(\Omega)$ represents the spherical harmonics and we can define $p_{\ell m}$ for the random field $p(\Omega)$ similarly. The inverse transform is then given by

$$\kappa_{\ell m} = \int d\Omega\, \kappa(\Omega) \mathcal{Y}_{\ell m}(\Omega), \tag{D2}$$

and the similar relation can be adopted for $p_{\ell m}$.

The effect of finite sky coverage for each field is characterized as

$$\tilde{\kappa}(\Omega) = W^\kappa(\Omega)\kappa(\Omega), \tag{D3}$$
$$\tilde{p}(\Omega) = W^p(\Omega)p(\Omega), \tag{D4}$$

where W^κ and W^p are the window function of sky masking for κ and p, respectively[6]. Thus, the harmonic modes in presence of masked region is expressed as

$$\tilde{X}_{\ell m} = \sum_{\ell'm'} W^X_{\ell m \ell' m'} X_{\ell'm'}, \tag{D5}$$

where $X = \kappa,\, p$ and $W^X_{\ell m \ell' m'}$ is defined by

$$W^X_{\ell m \ell' m'} = \int d\Omega\, \mathcal{Y}_{\ell'm'}(\Omega) W^X(\Omega) \mathcal{Y}^*_{\ell m}(\Omega). \tag{D6}$$

The estimators of power spectra on limited sky is defined by

$$\tilde{\boldsymbol{P}}(\ell) = \frac{1}{2\ell+1} \sum_m \langle \tilde{\boldsymbol{X}}_{\ell m} \tilde{\boldsymbol{X}}^\dagger_{\ell m} \rangle, \tag{D7}$$

where $\tilde{\boldsymbol{X}}_{\ell m} = (\tilde{\kappa}_{\ell m}, \tilde{p}_{\ell m})$ and $\langle \tilde{\boldsymbol{X}}_{\ell m} \tilde{\boldsymbol{X}}^\dagger_{\ell m} \rangle$ represents the following set of power spectra:

$$\langle \tilde{\boldsymbol{X}}_{\ell m} \tilde{\boldsymbol{X}}^\dagger_{\ell m} \rangle = \begin{pmatrix} \langle \tilde{\kappa}_{\ell m} \tilde{\kappa}^*_{\ell m} \rangle & \langle \tilde{\kappa}_{\ell m} \tilde{p}^*_{\ell m} \rangle \\ \langle \tilde{p}_{\ell m} \tilde{\kappa}^*_{\ell m} \rangle & \langle \tilde{p}_{\ell m} \tilde{p}^*_{\ell m} \rangle \end{pmatrix}. \tag{D8}$$

For the underlying field $\boldsymbol{X} = (\kappa, p)$, we can define the power spectra $\boldsymbol{P}(\ell)$ as

$$\boldsymbol{P}(\ell) = \langle \boldsymbol{X}_{\ell m} \boldsymbol{X}^\dagger_{\ell'm'} \rangle \delta_{\ell\ell'} \delta_{mm'}. \tag{D9}$$

Using Eqs. (D6) and (D9), we can find the relation between $\tilde{\boldsymbol{P}}(\ell)$ and $\boldsymbol{P}(\ell)$ as follows:

$$\tilde{\boldsymbol{P}}(\ell) = \frac{1}{2\ell+1} \sum_m \sum_{\ell'm'} \boldsymbol{W}_{\ell m \ell' m'} \boldsymbol{P}(\ell') (\boldsymbol{W}_{\ell m \ell' m'})^\dagger, \tag{D10}$$

$$= \sum_{\ell'} \boldsymbol{M}_{\ell\ell'} \boldsymbol{P}(\ell'), \tag{D11}$$

where $\boldsymbol{W}_{\ell m \ell' m'}$ is defined by

$$\boldsymbol{W}_{\ell m \ell' m'} = \begin{pmatrix} W^\kappa_{\ell m \ell' m'} & 0 \\ 0 & W^p_{\ell m \ell' m'} \end{pmatrix}. \tag{D12}$$

The matrix $\boldsymbol{M}_{\ell\ell'}$ represents the mode coupling effect due to masked region on power spectra, which is given by (in terms of $\boldsymbol{P}(\ell) = (P_{\kappa\kappa}(\ell), P_{p\kappa}(\ell), P_{pp}(\ell))^T$),

$$\boldsymbol{M}_{\ell\ell'} = \begin{pmatrix} M^{\kappa\kappa}_{\ell\ell'} & 0 & 0 \\ 0 & M^{p\kappa}_{\ell\ell'} & 0 \\ 0 & 0 & M^{pp}_{\ell\ell'} \end{pmatrix}, \tag{D13}$$

---

[6] In practice, W^p are not equal to W^κ. This is because peaks of convergence field are defined by that of *smoothed* convergence map. When the area with mask W^κ is smoothed, there would exist ill-defined pixels due to the convolution between W^κ and a filter function for smoothing. Therefore, we need to remove the ill-defined pixels to find peaks. This procedure makes the effective sky coverage of W^p smaller than that of W^κ.

where

$$M_{\ell\ell'}^{XY} = \sum_{mm'} W_{\ell m \ell' m'}^{X} \left[W_{\ell m \ell' m'}^{Y} \right]^{*}, \tag{D14}$$

$$= \frac{2\ell'+1}{4\pi} \sum_{L} \mathcal{W}_{L}^{XY} \begin{pmatrix} \ell & \ell' & L \\ 0 & 0 & 0 \end{pmatrix}^{2}, \tag{D15}$$

and $\mathcal{W}_{\ell}^{XY}$ is given by

$$\mathcal{W}_{\ell}^{XY} = \sum_{m} w_{\ell m}^{X} \left(w_{\ell m}^{Y} \right)^{*}, \tag{D16}$$

$$w_{\ell m}^{X} = \int d\Omega\, W^{X}(\Omega) \mathcal{Y}_{\ell m}(\Omega). \tag{D17}$$

Hence, we can construct the estimator of $\boldsymbol{P}(\ell)$ from $\tilde{\boldsymbol{P}}(\ell)$ as

$$\hat{\boldsymbol{P}}(\ell) = \sum_{\ell'} \boldsymbol{M}_{\ell\ell'}^{-1} \tilde{\boldsymbol{P}}(\ell'), \tag{D18}$$

where $\hat{\boldsymbol{P}}(\ell)$ is so-called psuedo-spectrum estimators.

Next, we consider the covariance of the pseudo-spectrum estimators. The covariance of $\hat{\boldsymbol{P}}(\ell)$ is defined by

$$\mathrm{Cov}[\hat{P}_{XY}(\ell), \hat{P}_{MN}(\ell')] = \langle \hat{P}_{XY}(\ell) \hat{P}_{MN}(\ell') \rangle - \langle \hat{P}_{XY}(\ell) \rangle \langle \hat{P}_{MN}(\ell) \rangle, \tag{D19}$$

where X, Y, M, N is set to be κ or p. When the underlying field follows non-Gaussian and there exist no masked regions, the covariance can be expressed as

$$\mathrm{Cov}[P_{XY}(\ell), P_{MN}(\ell')]_{\mathrm{all-sky}} = \frac{\delta_{\ell\ell'}}{2\ell+1} \left[P_{XM}(\ell) P_{YN}(\ell') + P_{XN}(\ell) P_{YM}(\ell') \right] + \frac{1}{2\ell+1} \frac{1}{2\ell'+1} \sum_{mm'} \langle X_{\ell m} Y_{\ell m}^{*} M_{\ell' m'} N_{\ell' m'}^{*} \rangle_{c}, \tag{D20}$$

where the first term of the right-hand side in Eq. (D20) represents the Gaussian contribution to the covariance matrix and the second term corresponds to the contribution of four-point correlation function due to non-Gaussianity in the underlying field. On the other hand, in presence of masked region, the covariance of the pseudo-spectrum estimators is expressed as (see also, e.g., Brown, Castro & Taylor (2005))

$$\mathrm{Cov}[\hat{P}_{XY}(\ell_1), \hat{P}_{MN}(\ell_2)] = \mathrm{Cov}[\hat{P}_{XY}(\ell_1), \hat{P}_{MN}(\ell_2)]_{\mathrm{NG}} + \sum_{\ell'_1 \ell'_2} (M^{XY})_{\ell_1 \ell'_1}^{-1} (M^{MN})_{\ell_2 \ell'_2}^{-1}$$

$$\times \sum_{\ell''_1 \ell''_2} \left[P_{AD}(\ell''_1) P_{BC}(\ell''_2) \mathcal{X}_{\ell'_1 \ell'_2 \ell''_1 \ell''_2}^{[XA, ND, MC, YB]} + P_{AC}(\ell''_1) P_{BD}(\ell''_2) \mathcal{X}_{\ell'_1 \ell'_2 \ell''_1 \ell''_2}^{[XA, MC, ND, YB]} \right], \tag{D21}$$

where all the possible combinations of A, B, C and D are taken into account in Eq. (D21) and $\mathcal{X}_{\ell\ell'\ell_1\ell_2}^{[XA, ND, MC, YB]}$ is given by

$$\mathcal{X}_{\ell\ell'\ell_1\ell_2}^{[XA, ND, MC, YB]} = \frac{1}{(2\ell+1)(2\ell'+1)} \sum W_{\ell m \ell_1 m_1}^{XA} \left(W_{\ell' m' \ell_1 m_1}^{ND} \right)^{*} W_{\ell' m' \ell_1 m_1}^{MC} \left(W_{\ell m \ell_2 m_2}^{YB} \right)^{*}. \tag{D22}$$

The summation in Eq. (D22) is taken over all m, m', m_1, m_2. Here, $W_{\ell m \ell' m'}^{XY}$ denotes $W^{p\kappa} = W^{\kappa p} = 0$, $W^{\kappa\kappa} = W^{\kappa}$, and $W^{pp} = W^{p}$ with Eq. (D6). The non-Gaussian term in Eq. (D21) is defined by

$$\mathrm{Cov}[\hat{P}_{XY}(\ell_1), \hat{P}_{MN}(\ell_2)]_{\mathrm{NG}} = \frac{1}{(2\ell_1+1)(2\ell_2+1)} \sum_{\ell'_1 \ell'_2} (M^{XY})_{\ell_1 \ell'_1}^{-1} (M^{MN})_{\ell_2 \ell'_2}^{-1}$$

$$\times \sum W_{\ell_1 m_1 \ell'_1 m'_1}^{XA} \left(W_{\ell_1 m_1 \ell''_1 m''_1}^{YB} \right)^{*} W_{\ell_2 m_2 \ell'_2 m'_2}^{MC} \left(W_{\ell_2 m_2 \ell''_2 m''_2}^{ND} \right)^{*} \langle A_{\ell'_1 m'_1} B_{\ell''_1 m''_1}^{*} C_{\ell'_2 m'_2} D_{\ell''_2 m''_2}^{*} \rangle_c \tag{D23}$$

where the second summation in Eq. (D23) is over all m_i, m'_i, m''_i, ℓ'_i, ℓ''_i $(i = 1, 2)$ and the values of A, B, C and D. Eqs. (D21) and (D23) clearly show that the complicated masked regions on sky would induce the *additional* mode-coupling of the covariance matrix of the pseudo-spectrum estimators.

REFERENCES

Allen S. W., Evrard A. E., Mantz A. B., 2011, Annual Review of Astronomy and Astrophysics, 49, 409

Anderson L. et al., 2014, Monthly Notices of the Royal Astronomical Society, 441, 24

Bartelmann M., Schneider P., 2001, Physics Reports, 340, 291

Becker M. R., 2013, Monthly Notices of the Royal Astronomical Society, 435, 115

Betoule M. et al., 2014, Astronomy and Astrophysics, 568, A22

Beutler F. et al., 2011, Monthly Notices of the Royal Astronomical Society, 416, 3017

Beutler F. et al., 2014, Monthly Notices of the Royal Astronomical Society, 443, 1065

Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, The Astrophysical Journal, 732, 122

Blake C. et al., 2011, Monthly Notices of the Royal Astronomical Society, 418, 1707

Brown M. L., Castro P. G., Taylor A. N., 2005, Monthly Notices of the Royal Astronomical Society, 360, 1262

Cooray A., Hu W., 2001, The Astrophysical Journal, 554, 56

Covone G., Sereno M., Kilbinger M., Cardone V. F., 2014, The Astrophysical Journal, 784, L25

Crocce M., Pueblas S., Scoccimarro R., 2006, Monthly Notices of the Royal Astronomical Society, 373, 369

Das S., Bode P., 2008, The Astrophysical Journal, 682, 1

Dietrich J. P., Hartlap J., 2010, Monthly Notices of the Royal Astronomical Society, 402, 1049

Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, Monthly Notices of the Royal Astronomical Society, 390, L64

Duffy A. R., Schaye J., Kay S. T., Vecchia C. D., Battye R. A., Booth C. M., 2010, Monthly Notices of the Royal Astronomical Society, 405, 2161

Efstathiou G., 2004, Monthly Notices of the Royal Astronomical Society, 349, 603

Fan Z., Shan H., Liu J., 2010, The Astrophysical Journal, 719, 1408

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, The Astrophysical Journal, 622, 759

Hamana T., Bartelmann M., Yoshida N., Pfrommer C., 2005, Monthly Notices of the Royal Astronomical Society, 356, 829

Hamana T., Colombi S. T., Thion A., Devriendt J. E. G. T., Mellier Y., Bernardeau F., 2002, Monthly Notices of the Royal Astronomical Society, 330, 365

Hamana T., Mellier Y., 2001, Monthly Notices of the Royal Astronomical Society, 176, 169

Hamana T., Oguri M., Shirasaki M., Sato M., 2012, Monthly Notices of the Royal Astronomical Society, 425, 2287

Hamana T., Takada M., Yoshida N., 2004, Monthly Notices of the Royal Astronomical Society, 350, 893

Hansen F. K., Górski K. M., 2003, Monthly Notices of the Royal Astronomical Society, 343, 559

Heitmann K., White M., Wagner C., Habib S., Higdon D., 2010, The Astrophysical Journal, 715, 104

Hennawi J. F., Spergel D. N., 2005, The Astrophysical Journal, 624, 59

Heymans C. et al., 2012, Monthly Notices of the Royal Astronomical Society, 427, 146

Hikage C., Takada M., Hamana T., Spergel D., 2011, Monthly Notices of the Royal Astronomical Society, 412, 65

Hilbert S., Hartlap J., White S. D. M., Schneider P., 2009, Astronomy and Astrophysics, 499, 31

Hilbert S., Marian L., Smith R. E., Desjacques V., 2012, Monthly Notices of the Royal Astronomical Society, 426, 2870

Hinshaw G. et al., 2013, The Astrophysical Journal Supplement Series, 208, 19

Hirata C. M., Seljak U., 2004, Physical Review D, 70, 063526

Hu W., 2001, Physical Review D, 64, 083005

Hu W., Kravtsov A. V., 2003, The Astrophysical Journal, 584, 702

Jain B., Seljak U., White S., 2000, The Astrophysical Journal, 530, 547

Jeong D., Komatsu E., Jain B., 2009, Physical Review D, 80, 123527

Kaiser N., 1992, The Astrophysical Journal, 388, 272

Kayo I., Takada M., Jain B., 2013, Monthly Notices of the Royal Astronomical Society, 429, 344

Kilbinger M. et al., 2013, Monthly Notices of the Royal Astronomical Society, 430, 2200

Kratochvil J. M., Haiman Z., May M., 2010, Physical Review D, 81, 043519

Lewis A., Challinor A., Lasenby A., 2000, The Astrophysical Journal, 538, 473

Lilje P. B., 1992, The Astrophysical Journal, 386, L33

Limber D. N., 1954, The Astrophysical Journal, 119, 655

Liu J., Haiman Z., Hui L., Kratochvil J. M., May M., 2014a, Physical Review D, 89, 023515

Liu X., Wang Q., Pan C., Fan Z., 2014b, The Astrophysical Journal, 784, 31

Marian L., Smith R. E., Hilbert S., Schneider P., 2012, Monthly Notices of the Royal Astronomical Society, 423, 1711

Marian L., Smith R. E., Hilbert S., Schneider P., 2013, Monthly Notices of the Royal Astronomical Society, 432, 1338

Maturi M., Angrick C., Pace F., Bartelmann M., 2010, Astronomy and Astrophysics, 519, A23

Maturi M., Meneghetti M., Bartelmann M., Dolag K., Moscardini L., 2005, Astronomy and Astrophysics, 442, 851

Miyazaki S., Hamana T., Ellis R. S., Kashikawa N., Massey R. J., Taylor J., Refregier A., 2007, The Astrophysical Journal,

669, 714

Mohammed I., Martizzi D., Teyssier R., Amara A., 2014, eprint arXiv:1410.6826

Munshi D., Valageas P., Vanwaerbeke L., Heavens a., 2008, Physics Reports, 462, 67

Navarro J., Frenk C., White S., 1997, The Astrophysical Journal, 490, 493

Nishimichi T. et al., 2009, Publications of the Astronomical Society of Japan, 61, 321

Oguri M., Bayliss M. B., Dahle H., Sharon K., Gladders M. D., Natarajan P., Hennawi J. F., Koester B. P., 2012, Monthly Notices of the Royal Astronomical Society, 420, 3213

Oguri M., Takada M., 2011, Physical Review D, 83, 023008

Okabe N., Smith G. P., Umetsu K., Takada M., Futamase T., 2013, The Astrophysical Journal Letters, 769, L35

Osato K., Shirasaki M., Yoshida N., 2015, ArXiv e-prints

Peacock J. A., Dodds S. J., 1996, Monthly Notices of the Royal Astronomical Society, 280, L19

Planck Collaboration et al., 2014, Astronomy and Astrophysics, 571, A16

Sánchez A. G. et al., 2012, Monthly Notices of the Royal Astronomical Society, 425, 415

Sato M., Hamana T., Takahashi R., Takada M., Yoshida N., Matsubara T., Sugiyama N., 2009, The Astrophysical Journal, 701, 945

Schirmer M., Erben T., Hetterscheidt M., Schneider P., 2007, Astronomy and Astrophysics, 462, 875

Schmidt F., Rozo E., 2011, The Astrophysical Journal, 735, 119

Schneider P., 1996, Monthly Notices of the Royal Astronomical Society, 283, 837

Semboloni E., Hoekstra H., Schaye J., 2013, Monthly Notices of the Royal Astronomical Society, 434, 148

Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, Monthly Notices of the Royal Astronomical Society, 417, 2020

Shan H. et al., 2012, The Astrophysical Journal, 748, 56

Shirasaki M., Yoshida N., 2014, The Astrophysical Journal, 786, 43

Smith R. E. et al., 2003, Monthly Notices of the Royal Astronomical Society, 341, 1311

Springel V., 2005, Monthly Notices of the Royal Astronomical Society, 364, 1105

Takada M., Bridle S., 2007, New Journal of Physics, 9, 446

Takada M., Jain B., 2004, Monthly Notices of the Royal Astronomical Society, 348, 897

Takada M., Spergel D. N., 2014, Monthly Notices of the Royal Astronomical Society, 441, 2456

Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, The Astrophysical Journal, 761, 152

Teyssier R. et al., 2009, Astronomy and Astrophysics, 497, 335

Valageas P., Nishimichi T., 2011, Astronomy and Astrophysics, 527, A87

van Waerbeke L., 2000, Monthly Notices of the Royal Astronomical Society, 313, 524

Yang X., Kratochvil J., Huffenberger K., Haiman Z., May M., 2013, Physical Review D, 87, 023511

Yang X., Kratochvil J. M., Wang S., Lim E. A., Haiman Z., May M., 2011, Physical Review D, 84, 043529

Zentner A. R., Semboloni E., Dodelson S., Eifler T., Krause E., Hearin A. P., 2013, Physical Review D, 87, 043509